

Università degli Studi di Padova
Dipartimento di Scienze Statistiche
Corso di Laurea Magistrale in
Scienze Statistiche



**Stima e previsione per modelli INAR(1) con
distribuzione dell'errore Binomiale e Binomiale
Negativa**

Relatore Prof. Luisa Bisaglia
Dipartimento di Scienze Statistiche

Laureando: Paolo Gorgi
Matricola N. 1035015

Anno Accademico 2012/2013

Indice

Introduzione	2
1 Il modello PoINAR(1)	5
1.1 Processi INAR(1)	5
1.2 Generalità	7
1.3 Stima dei parametri	8
1.4 Previsioni	9
2 INAR(1) non-Poisson	13
2.1 Motivazioni	13
2.2 Innovazione Binomiale	14
2.2.1 Generalità	14
2.2.2 Procedura di stima	15
2.2.3 Previsioni	17
2.3 Innovazione Binomiale Negativa	21
2.3.1 Generalità	21
2.3.2 Procedura di stima	22
2.3.3 Previsioni	24
3 Esperimento Monte Carlo	29
3.1 Informazioni generali	29
3.2 Stimatori di massima verosimiglianza	30
3.2.1 Motivazioni e obbiettivi	30

3.2.2	Innovazione Poisson	31
3.2.3	Innovazione Binomiale	36
3.2.4	Innovazione Binomiale Negativa	42
3.3	Previsioni	48
3.3.1	Obbiettivi e criteri utilizzati	48
3.3.2	Innovazione Poisson	50
3.3.3	Innovazione Binomiale	54
3.3.4	Innovazione Binomiale negativa	59
4	Applicazione su dati reali	65
4.1	La serie storica	65
4.2	Confronto tra modelli	68
	Conclusioni	73
	Bibliografia	75

Introduzione

Serie storiche di dati a valori interi possono essere incontrate in diversi contesti e applicazioni, basti pensare, ad esempio, al numero di ospiti di un Hotel in un dato giorno, al numero di pazienti in una lista d'attesa, al numero di persone affette da una certa malattia in un dato mese, al numero di incidenti sul lavoro in un'azienda, etc. Questa tipologia di serie, per poter essere modellata in modo adeguato, richiede l'utilizzo di appositi modelli che tengano conto della caratteristica di conteggio delle osservazioni. In letteratura sono presenti diverse soluzioni a riguardo; per una rassegna si può fare riferimento a Fokianos (2012).

In questa tesi è trattata la classe di modelli INAR(1) (*first-order INteger-valued AutoRegressive*) proposti da McKenzie (1985) e Al-Osh e Alzaid (1987).

Solitamente, uno dei principali obiettivi della modellazione di una serie storica è quello di ottenere delle previsioni per le osservazioni future. Nei modelli INAR(1) la previsione presenta alcune complicazioni rispetto ai modelli per serie continue e per questo motivo è un argomento particolarmente trattato in letteratura. Utilizzando le classiche tecniche di previsione basate sul valore atteso condizionale, difficilmente si ottengono delle previsioni intere e quindi coerenti con la caratteristica di conteggio della serie. Relativamente a questo problema, è stata proposta come soluzione l'utilizzo della mediana condizionale della distribuzione di probabilità h passi in avanti, Freeland e McCabe (2004). Questa soluzione è stata applicata per i modelli PoINAR(1), cioè modelli INAR(1) con il termine d'errore del processo distribuito come una

Poisson. Teoricamente è comunque possibile adottare una qualsiasi distribuzione discreta e implementare la procedura previsiva basata sulla distribuzione di probabilità h passi in avanti, McCabe e Martin (2005).

La distribuzione Poisson ha il grosso limite di imporre l'equidispersione del processo (media pari alla varianza) e questo, in molti casi, è poco plausibile. Delle possibili alternative sono la distribuzione Binomiale che impone sottodispersione (media maggiore della varianza) e la distribuzione Binomiale Negativa che impone sovradispersione (media minore della varianza).

Lo scopo principale di questa tesi è quello di ricavare le procedure di stima e previsione per i modelli INAR(1) con distribuzione dell'errore Binomiale e Binomiale Negativa, in modo da poter verificare se queste possono essere considerate delle valide alternative alla distribuzione Poisson. La verifica è fatta attraverso un confronto sulle *performance* previsive dei modelli in differenti circostanze.

La tesi è strutturata in quattro capitoli. Nel primo capitolo sono presentati i modelli INAR(1) e sono esposte le procedure di stima e previsione per il caso Poisson. Nel secondo capitolo vengono sviluppate le procedure di stima e previsione per il modello con la Binomiale e la Binomiale negativa. Il terzo capitolo è suddiviso in due parti: nella prima sono presentati i risultati di uno studio di simulazione sul comportamento degli stimatori di massima verosimiglianza, mentre nella seconda sono esposti i risultati delle simulazioni riguardanti il confronto tra i diversi modelli in fase previsiva. Infine, nel quarto capitolo, è presentata un'applicazione su dati reali, in cui è fatto un raffronto tra l'utilizzo della distribuzione di Poisson e la distribuzione Binomiale Negativa.

Capitolo 1

Il modello PoINAR(1)

1.1 Processi INAR(1)

Nel trattare serie storiche per dati di conteggio risulta sicuramente apprezzabile specificare dei modelli i quali considerino il fatto che le osservazioni sono di tipo discreto. I modelli della classe INAR (*INteger-valued AutoRegressive*) tengono conto di questa peculiarità delle serie di conteggio attraverso la definizione dell'operatore *thinning*, Stuetal e Van Harn (1979). Data una variabile aleatoria discreta non negativa X e un parametro $\alpha \in [0, 1]$; l'operatore *thinning* \circ è definito come

$$\alpha \circ X = \sum_{i=1}^X Y_i$$

dove Y_i è una successione i.i.d di variabili aleatorie Bernulliane con parametro α . Questo implica che la distribuzione di $\alpha \circ X$ condizionatamente a $X = x$ è Binomiale di parametri x e α . In generale è comunque possibile utilizzare anche distribuzioni diverse dalla Bernulli. Una volta definito l'operatore *thinning*, il processo INAR(1) $\{X_t : t = 0, 1, 2, \dots\}$ può essere definito come:

$$X_t = \alpha \circ X_{t-1} + \epsilon_t \tag{1.1}$$

dove $\alpha \in [0, 1]$ e $\{\epsilon_t\}$ è una successione i.i.d di variabili aleatorie discrete non negative con media μ e varianza σ^2 . La condizione che $0 \leq \alpha < 1$ garantisce la stazionarietà del processo, Grunwald et al. (2000).

La classe di modelli INAR(1) consente di ottenere una semplice e intuitiva interpretazione del processo generatore dei dati; infatti, il valore di conteggio all'istante t può essere visto come la somma tra i sopravvissuti dal tempo $t - 1$ e i nuovi nati al tempo t . I sopravvissuti sono la realizzazione della variabile $\alpha \circ X_{t-1}$ e i nuovi nati sono la realizzazione dell'innovazione ϵ_t . È anche possibile estendere i modelli INAR considerando più ritardi nell'equazione del processo (INAR(p)), Al-Osh e Alzaid (1990) e Jayakumar (1995), ma facendo questo si perde la facile interpretabilità che è uno dei punti di forza dei modelli INAR(1).

Il processo (1.1) può anche essere scritto nella forma

$$X_t = \sum_{i=0}^{+\infty} \alpha^i \circ \epsilon_{t-i} \quad (1.2)$$

la quale è ricavata considerando che $\alpha \circ (\alpha \circ X) = \alpha^2 \circ X$, Al-Osh e Alzaid (1987).

La media e la varianza marginale del processo sono date, rispettivamente, da:

$$\mu_x = E(X_t) = \frac{\mu}{1 - \alpha} \quad (1.3)$$

$$\sigma_x^2 = Var(X_t) = \frac{\alpha\mu + \sigma^2}{1 - \alpha^2} \quad (1.4)$$

mentre, la funzione di autocovarianza è pari a

$$\gamma(h) = Cov(X_t, X_{t-h}) = \alpha^h \gamma(0) = \alpha^h \sigma_x^2 \quad (1.5)$$

Dalla (1.4) e (1.5) è facile ricavare la funzione di autocorrelazione del processo

$$\rho(h) = \text{Corr}(X_t, X_{t-h}) = \alpha^h \quad (1.6)$$

La (1.6) mostra come la funzione di autocorrelazione dei processi INAR(1) sia equivalente a quella di un classico AR(1), con la correlazione che decresce esponenzialmente all'aumentare del ritardo temporale h . Chiaramente per gli INAR(1), visto che $\alpha \in [0, 1)$, $\rho(h)$ è sempre positiva.

La distribuzione marginale del processo INAR(1) dipende dalla specificazione della distribuzione del termine ϵ_t ; in particolare, data una distribuzione del termine d'errore non è scontato che la marginale appartenga alla stessa famiglia di distribuzioni.

1.2 Generalità

La definizione (1.1) ammette che la distribuzione del termine di innovazione ϵ_t possa essere una qualunque variabile di conteggio con media e varianza finite. In letteratura sostanzialmente sono trattate due diverse specificazioni; la prima considera come marginale del processo la distribuzione Binomiale Negativa, McKenzie (1986), mentre la seconda assume che il processo abbia distribuzione marginale di Poisson. La specificazione con marginale Binomiale Negativa, probabilmente a causa di alcune complicazioni legate alla distribuzione di ϵ_t , non riscontra molte applicazioni. La specificazione di Poisson è sicuramente quella più trattata in letteratura e su di essa sono state sviluppate le procedure basate sulla distribuzione di probabilità h passi in avanti per ottenere previsioni coerenti, Freeland e McCabe (2004). La vasta applicazione della Poisson è probabilmente legata alle conseguenti proprietà sulla distribuzione condizionale e sulla distribuzione marginale del processo. Infatti, se ϵ_t si

distribuisce come una Poisson di parametro λ , allora la distribuzione marginale del processo è una Poisson di parametro $\lambda/(1 - \alpha)$. I processi INAR(1) con distribuzione di Poisson sono detti PoINAR(1).

Nel seguito di questo capitolo saranno trattate le tematiche relative alla stima e alla previsione con i modelli PoINAR(1).

1.3 Stima dei parametri

La stima dei parametri α e λ del modello PoINAR(1) può essere ottenuta attraverso diverse procedure. In questa tesi sarà considerato il metodo della massima verosimiglianza dato che questo sembra essere preferibile sia rispetto allo stimatore di Yule-walker sia rispetto allo stimatore dei Minimi Quadrati Condizionati, Al-Osh e Alzaid (1987).

Dato un campione $x = \{x_0, x_1, \dots, x_n\}$ la funzione di verosimiglianza è

$$L(\alpha, \lambda|x) = \left(\prod_{t=1}^n p(X_t|X_{t-1}) \right) p(X_0) \quad (1.7)$$

dove la distribuzione condizionale di X_t dato $X_{t-1} = x_{t-1}$ è la convoluzione di una distribuzione di Poisson di parametro λ con una Binomiale di parametri x_{t-1} e α .

La distribuzione di $X_t|X_{t-1}$ risulta

$$\begin{aligned} p(X_t|X_{t-1}) &= Pr\{X_t = x_t|X_{t-1} = x_{t-1}\} = \\ &= \exp(-\lambda) \sum_{i=0}^{\min(x_{t-1}, x_t)} \frac{\lambda^{x_t-i}}{(x_t-i)!} \binom{x_{t-1}}{i} \alpha^i (1-\alpha)^{x_{t-1}-i} \end{aligned} \quad (1.8)$$

con $x_t \in \{0, 1, 2, \dots\}$.

Per quanto riguarda la distribuzione marginale di X_0 nella (1.7), questa può

essere esplicitata dato che è una Poisson di parametro $\lambda/(1 - \alpha)$, oppure, in maniera alternativa, può essere esclusa ottenendo così la funzione di verosimiglianza condizionata a $X_0 = x_0$.

Le stime di massima verosimiglianza si ottengono massimizzando la (1.7) rispetto a α e λ . Questo problema di massimizzazione, per essere risolto, necessita di procedere per via numerica dato che non sono state trovate soluzioni esplicite.

Le stime così ottenute godono delle usuali proprietà asintotiche degli stimatori di massima verosimiglianza:

$$\sqrt{n}(\hat{\alpha}_n, \hat{\lambda}_n)' \overset{a}{\sim} N((\alpha_0, \lambda_0)', V) \quad (1.9)$$

dove V è la matrice di varianza e covarianza che a sua volta può essere stimata per via numerica.

1.4 Previsioni

Nell'ambito delle serie storiche per dati di conteggio la previsione è un argomento di particolare interesse dato che, per ottenere risultati coerenti, anche i valori previsti devono essere discreti. Nonostante i modelli INAR(1) siano ben specificati per considerare la caratteristica di conteggio del processo, se si applica la classica tecnica di previsione basata sul minimizzare l'errore quadratico medio di previsione, utilizzando quindi il valore atteso condizionale h passi in avanti $E[X_{t+h}|X_t = x_t]$ come previsore, difficilmente si ottengono previsioni discrete. A tale scopo una possibilità è quella di utilizzare il valore che minimizza l'errore assoluto medio di previsione h passi in avanti, questo valore è la mediana della distribuzione condizionale $p(X_{t+h}|X_t)$. Questo approccio, relativamente ai modelli PoINAR(1), è stato applicato in diversi lavori, fra cui, ad esempio, Freeland e McCabe (2004) e Silva et al. (2009), e richiede

il calcolo dell'intera distribuzione condizionale ad h passi da cui poi ricavare la mediana. Possibili alternative sono l'utilizzo della moda di $p(X_{t+h}|X_t)$, oppure, sebbene sia una soluzione poco elegante, si può arrotondare la media condizionale all'intero più vicino e ottenere così una previsione intera.

Per ricavare la $p(X_{t+h}|X_t)$ si scrive il processo nella forma

$$X_{t+h} = \alpha^h \circ X_t + \sum_{i=0}^{h-1} \alpha^i \circ \epsilon_{t+h-i} \quad (1.10)$$

dalla (1.10) si può vedere come la distribuzione condizionale di X_{t+h} dato $X_t = x_t$ sia la convoluzione tra la distribuzione di $(\alpha_h \circ X_t)|X_t = x_t$, che è una Binomiale di parametri x_t e α , e la distribuzione di $\sum_{i=0}^{h-1} \alpha^i \circ \epsilon_{t+h-i}$. Si dimostra, Freeland (1998), che la distribuzione di $\sum_{i=0}^{h-1} \alpha^i \circ \epsilon_{t+h-i}$ è una Poisson di parametro $\lambda(1 - \alpha^h)/(1 - \alpha)$ dato che ogni elemento $\alpha^i \circ \epsilon_{t+h-i}$ della sommatoria $\sum_{i=0}^{h-1} \alpha^i \circ \epsilon_{t+h-i}$ è Poisson. Questo risultato può anche essere dimostrato ricavando la funzione generatrice dei momenti di $\alpha^i \circ \epsilon_{t+h-i}$

$$\begin{aligned} M_{\alpha^i \circ \epsilon_{t-i}}(s) &= E\left[\exp\{s(\alpha^i \circ \epsilon_{t-i})\}\right] = \\ &= E\left[E[\exp\{s(\alpha^i \circ \epsilon_{t-i})\}|\epsilon_{t-i}]\right] = \\ &= E\left[(\alpha^i \exp\{s\} + (1 - \alpha^i))^{\epsilon_{t-i}}\right] = \\ &= E\left[\exp\left\{\epsilon_{t-i} \log\left(\alpha^i \exp\{s\} + (1 - \alpha^i)\right)\right\}\right] = \\ &= E\left[\exp\{s' \epsilon_{t-i}\}\right] = \\ &= \exp\left\{\lambda\left(\exp\{s'\} - 1\right)\right\} = \\ &= \exp\left\{\lambda\left(\exp\left\{\log\left(\alpha^i \exp\{s\} + (1 - \alpha^i)\right)\right\} - 1\right)\right\} = \\ &= \exp\left\{\lambda\alpha^i\left(\exp\{s\} - 1\right)\right\} \end{aligned} \quad (1.11)$$

con $s' = \log\left(\alpha^i \exp\{s\} + (1 - \alpha^i)\right)$.

Dalla (1.11) si vede come la funzione generatrice dei momenti di $\alpha^i \circ \epsilon_{t+h-i}$ sia quella di una Poisson di parametro $\alpha^i \lambda$.

La funzione di probabilità di $X_{t+h}|X_t = x_t$ è pari quindi a:

$$\begin{aligned}
p(X_{t+h}|X_t) &= Pr\{X_{t+h} = x_{t+h}|X_t = x_t\} = \\
&= \exp\left(-\lambda \frac{1-\alpha^h}{1-\alpha}\right) \sum_{i=0}^{Min(x_{t+h}, x_t)} \frac{1}{(x_{t+h} - i)!} \\
&\quad \times \left(\lambda \frac{1-\alpha^h}{1-\alpha}\right)^{x_{t+h}-i} \binom{x_t}{i} (\alpha^h)^i (1-\alpha^h)^{x_t-i} \quad (1.12)
\end{aligned}$$

con $x_{t+h} \in \{0, 1, 2, \dots\}$ e $h = 1, 2, 3, \dots$

A partire dalla (1.12) e sostituendo ai parametri le loro stime si ottiene una previsione della funzione di probabilità di X_t h passi in avanti. Calcolando la mediana o la moda di questa si possono ottenere previsioni h passi in avanti coerenti con la caratteristica di conteggio del processo. Si può, inoltre, notare come, per $h \rightarrow \infty$, la (1.12) converga alla distribuzione di probabilità di una Poisson di parametro $\lambda/(1-\alpha)$, che è la marginale del processo.

Capitolo 2

INAR(1) non-Poisson

2.1 Motivazioni

Nel precedente capitolo si è visto come per il modello PoINAR(1) esistano alcune proprietà che ne facilitano l'applicazione. Infatti, per questo modello, la distribuzione marginale del processo è nota ed è, inoltre, possibile ricavare una forma esplicita della distribuzione condizionale di previsione ad h passi.

Il modello PoINAR(1) presenta però anche alcune limitazioni a causa dell'assunzione di equidispersione (media pari alla varianza del processo) e alla bassa flessibilità della distribuzione di Poisson che ha un solo parametro. Per questo motivo può essere utile considerare differenti specificazioni del termine di errore in modo da garantire una maggiore flessibilità attraverso l'uso di distribuzioni con più parametri. Le due distribuzioni trattate nel seguito sono la distribuzione Binomiale che assume sottodispersione e la distribuzione Binomiale Negativa che assume sovradispersione.

In questo capitolo sono sviluppate le procedure di stima e previsione per i modelli INAR(1) assumendo che il termine ϵ_t nella (1.1) sia distribuito come una variabile aleatoria Binomiale e come una variabile aleatoria Binomiale Negativa. Essendo lo scopo di questo lavoro quello di cercare delle valide alternative al modello PoINAR(1), nel terzo capitolo, queste procedure saranno

testate per verificare la loro bontà da un punto di vista previsivo.

2.2 Innovazione Binomiale

2.2.1 Generalità

In questa sezione si assume che l'innovazione del processo sia distribuita come una variabile aleatoria Binomiale di parametri m e p . Considerando il valore di conteggio assunto da X_t come la somma tra i sopravvissuti dall'istante $t - 1$ e i nuovi nati all'istante t , si ha che i nuovi nati sono pari al numero di successi su m prove, dove ogni prova ha probabilità p di successo. Una delle differenze che si evidenziano rispetto al caso Poisson è che, conoscendo il conteggio all'istante $t - 1$, il supporto del conteggio all'istante t è limitato. La variabile aleatoria $X_t|X_{t-1} = x_{t-1}$ ha supporto pari a $\{0, 1, 2, \dots, x_{t-1} + m\}$.

La Binomiale ha la caratteristica di avere la media maggiore della varianza e questa può essere considerata un'alternativa rispetto alla media pari alla varianza della distribuzione di Poisson. La distribuzione marginale del processo non è una distribuzione nota, ma possono comunque essere calcolati i primi due momenti. Sfruttando i risultati (1.3), (1.4) e (1.5) si ottiene che: la media marginale è pari a

$$\mu_x = E(X_t) = \frac{mp}{1 - \alpha} \quad (2.1)$$

la varianza marginale risulta

$$\sigma_x^2 = Var(X_t) = \frac{\alpha mp + mp(1 - p)}{1 - \alpha^2} \quad (2.2)$$

mentre, la funzione di autocovarianza è

$$\gamma(k) = Cov(X_t, X_{t-k}) = \alpha^k \gamma(0) = \alpha^k \frac{\alpha mp + mp(1-p)}{1 - \alpha^2} \quad (2.3)$$

Dalla (2.1) e dalla (2.2) è facile vedere come, oltre all'innovazione, anche la marginale del processo abbia la media maggiore della varianza.

2.2.2 Procedura di stima

Per la stima dei parametri con il metodo della massima verosimiglianza si deve tener presente che la distribuzione Binomiale, quando il parametro m non è noto, non appartiene alla famiglia esponenziale e non possiede verosimiglianza regolare dato che il supporto dipende dai parametri. Per questo motivo potrebbero non essere garantite le usuali proprietà degli stimatori di massima verosimiglianza.

I parametri del modello da stimare sono α , m e p , e la funzione di verosimiglianza è data da

$$L(\alpha, m, p|x) = \left(\prod_{t=1}^n p(X_t|X_{t-1}) \right) p(X_0) \quad (2.4)$$

essendo X_t pari alla somma di due variabili aleatorie indipendenti, la distribuzione di $X_t|X_{t-1} = x_{t-1}$ è la convoluzione delle distribuzioni di probabilità di $\alpha \circ x_{t-1}$ e ϵ_t

$$\begin{aligned} p(X_t|X_{t-1}) &= Pr\{X_t = x_t|X_{t-1} = x_{t-1}\} = \\ &= Pr\{\alpha \circ X_{t-1} + \epsilon_t = x_t|X_{t-1} = x_{t-1}\} = \\ &= \sum_{i=-\infty}^{+\infty} Pr\{\alpha \circ x_{t-1} = i\} Pr\{\epsilon_t = x_t - i\} = \\ &= \sum_{i=\max(0, x_t-m)}^{\min(x_t, x_{t-1})} Pr\{\alpha \circ x_{t-1} = i\} Pr\{\epsilon_t = x_t - i\} \end{aligned}$$

(2.5)

nella sommatoria della (2.5) si ha che i varia dal $\max(0, x_t - m)$ al $\min(x_t, x_{t-1})$ considerando che $Pr\{\alpha \circ x_{t-1} = i\}$ è non nulla per $i \in \{0, 1, 2, \dots, x_{t-1}\}$, e $Pr\{\epsilon_t = x_t - i\}$ è non nulla per $x_t - i \in \{0, 1, 2, \dots, m\}$.

Sostituendo nella (2.5) per ogni distribuzione di probabilità la relativa espressione, si ottiene che

$$\begin{aligned}
p(X_t|X_{t-1}) &= \\
&= \sum_{i=\max(0, x_t-m)}^{\min(x_t, x_{t-1})} \binom{x_{t-1}}{i} \alpha^i (1-\alpha)^{x_{t-1}-i} \binom{m}{x_t-i} p^{x_t-i} (1-p)^{m-x_t+i} \\
&= (1-\alpha)^{x_{t-1}} (1-p)^m \left(\frac{p}{(1-p)} \right)^{x_t} \sum_{i=\max(0, x_t-m)}^{\min(x_t, x_{t-1})} \binom{x_{t-1}}{i} \\
&\quad \times \binom{m}{x_t-i} \left(\frac{\alpha(1-p)}{p(1-\alpha)} \right)^i
\end{aligned} \tag{2.6}$$

per $x_t \in \{0, 1, 2, \dots, (x_{t-1} + m)\}$.

A differenza del caso Poisson, la funzione di probabilità marginale del processo non è nota e la distribuzione di X_0 nella (2.4) non può essere ottenuta in maniera esplicita. Si utilizza quindi la funzione di verosimiglianza condizionata a $X_0 = x_0$ che è: $L(\alpha, m, p|x_0) = \prod_{t=1}^n p(X_t|X_{t-1})$, da cui, sostituendo l'espressione (2.6) e prendendo il logaritmo, si ricava la funzione di log-verosimiglianza condizionata

$$\begin{aligned}
l(\alpha, m, p|x_0) &= \log(L(\alpha, m, p|x_0)) = \sum_{t=1}^n \log(p(X_t|X_{t-1})) = \\
&= \left(\sum_{t=1}^n x_{t-1} \right) \log(1-\alpha) + mn \log(1-p) + \left(\sum_{t=1}^n x_t \right) \log \left(\frac{p}{(1-p)} \right) + \\
&\quad + \sum_{t=1}^n \log \left(\sum_{i=\max(0, x_t-m)}^{\min(x_t, x_{t-1})} \binom{x_{t-1}}{i} \binom{m}{x_t-i} \left(\frac{\alpha(1-p)}{p(1-\alpha)} \right)^i \right)
\end{aligned} \tag{2.7}$$

Le stime di massima verosimiglianza sono i valori di α , m e p che massimizzano la (2.7). Per questo problema di massimo non è calcolabile una soluzione esplicita e si deve quindi procedere per via numerica attraverso un algoritmo di ottimizzazione. La massimizzazione della log-verosimiglianza deve essere fatta tenendo conto dei valori che possono essere assunti dai parametri; infatti $\alpha \in [0, 1)$, $m \in \{1, 2, \dots\}$ e $p \in (0, 1)$. In questa circostanza sorgono alcune problematiche dato che il parametro m può assumere soltanto valori discreti e gli algoritmi di ottimizzazione generalmente funzionano soltanto per valori continui. Una possibile soluzione può essere data dalla seguente procedura:

1. Fissare l'insieme di valori $\{m_{min}, m_{min} + 1, m_{min} + 2, \dots, m_{max}\}$ che possono essere assunti dal parametro m . Dove: m_{min} può essere posto pari al $\max(\Delta x_1, \Delta x_2, \dots, \Delta x_n)$, con $\Delta x_t = x_t - x_{t-1}$ e m_{max} può essere fissato arbitrariamente a patto che sia sufficientemente elevato.
2. Per ognuno degli m fissati, attraverso un ottimizzatore numerico, si ricavano i valori dei parametri α e p che rendono massima la funzione di log-verosimiglianza. In questo modo si ottengono, per ogni m fissato, le stime di α , p e il valore della log-verosimiglianza.
3. Le stime di massima verosimiglianza sono i valori di α , m e p per cui il valore della log-verosimiglianza, calcolato fissando m , è massimo.

La procedura appena esposta ha lo svantaggio di essere computazionalmente pesante e l'utilizzo di un algoritmo che consenta di ottimizzare direttamente rispetto a valori discreti è sicuramente preferibile da questo punto di vista.

2.2.3 Previsioni

Come già visto nel primo capitolo, per ottenere delle previsioni coerenti con la caratteristica di conteggio dei dati, si deve prevedere l'intera distribuzione

di probabilità h passi in avanti. Quando l'innovazione è Binomiale non è possibile ricavare una forma generica della funzione di probabilità ad h passi, come avviene invece per i modelli PoINAR(1) (1.12). Questo è dovuto al fatto che nella (1.10) la distribuzione di $\sum_{i=0}^{h-1} \alpha^i \circ \epsilon_{t+h-i}$ non ha una formulazione esplicita per un generico valore di h ; infatti ogni elemento $\alpha^i \circ \epsilon_{t+h-i}$ della sommatoria è distribuito come una Binomiale con differente probabilità di successo, come si può vedere sfruttando la funzione generatrice dei momenti:

$$\begin{aligned}
M_{\alpha^i \circ \epsilon_{t-i}}(s) &= E\left[\exp\{s(\alpha^i \circ \epsilon_{t-i})\}\right] = \\
&= E\left[E[\exp\{s(\alpha^i \circ \epsilon_{t-i})\} | \epsilon_{t-i}]\right] = \\
&= E\left[(\alpha^i \exp\{s\} + (1 - \alpha^i))^{\epsilon_{t-i}}\right] = \\
&= E\left[\exp\left\{\epsilon_{t-i} \log\left(\alpha^i \exp\{s\} + (1 - \alpha^i)\right)\right\}\right] = \\
&= E\left[\exp\{s' \epsilon_{t-i}\}\right] = \\
&= \left(p \exp\{s'\} + (1 - p)\right)^m = \\
&= \left(p \exp\left\{\log\left(\alpha^i \exp\{s\} + (1 - \alpha^i)\right)\right\} + (1 - p)\right)^m = \\
&= \left(p(\alpha^i \exp\{s\} + (1 - \alpha^i)) + (1 - p)\right)^m = \\
&= \left(p\alpha^i \exp\{s\} + (1 - p\alpha^i)\right)^m \tag{2.8}
\end{aligned}$$

con $s' = \log(\alpha^i \exp\{s\} + (1 - \alpha^i))$.

Dalla (2.8) si deduce che ogni elemento $\alpha^i \circ \epsilon_{t+h-i}$ ha distribuzione Binomiale di parametri m e $\alpha^i p$, e quindi la sommatoria $\sum_{i=0}^{h-1} \alpha^i \circ \epsilon_{t+h-i}$, oltre a non avere distribuzione Binomiale, non ha una distribuzione esplicitabile per un generico h . Si può vedere come la distribuzione di $X_{t+h} | X_t = x_t$ sia la distribuzione della somma di $h + 1$ Binomiali con differente parametro di probabilità di successo.

Nonostante non sia possibile avere una formula generale per la distribuzione

condizionale si può comunque procedere calcolando la funzione di probabilità h passi in avanti per alcuni valori di h . Nel seguito sono ottenute le probabilità condizionali per $h = 1$ e $h = 2$.

La distribuzione per $h = 1$ di $X_{t+h}|X_t = x_t$ può essere facilmente ricavata sfruttando il risultato (2.6)

$$\begin{aligned}
p(X_{t+1}|X_t) &= Pr(X_{t+1} = x_{t+1}|X_t = x_t) = \\
&= (1 - \alpha)^{x_t} (1 - p)^m \left(\frac{p}{(1 - p)} \right)^{x_{t+1}} \\
&\quad \times \sum_{i=\max(0, x_t - m)}^{\min(x_{t+1}, x_t)} \binom{x_t}{i} \binom{m}{x_{t+1} - i} \left(\frac{\alpha(1 - p)}{p(1 - \alpha)} \right)^i \quad (2.9)
\end{aligned}$$

per $x_{t+1} \in \{0, 1, 2, \dots, (x_t + m)\}$.

Per $h = 2$, in riferimento alla (1.10), $X_{t+2}|X_t = x_t$ è la somma di $\alpha^2 \circ x_t$, $\alpha \circ \epsilon_{t+1}$ e ϵ_{t+2} ; queste tre variabili aleatorie sono indipendenti e distribuite, rispettivamente, come una Binomiale di parametri x_t e α^2 , una Binomiale di parametri m e αp , e una Binomiale di parametri m e p .

La $p(X_{t+2}|X_t)$ è pari alla la convoluzione tra la distribuzione di $\alpha^2 \circ x_t$ e la distribuzione di $\alpha \circ \epsilon_{t+1} + \epsilon_{t+2}$

$$\begin{aligned}
p(X_{t+2}|X_t) &= Pr(X_{t+2} = x_{t+2}|X_t = x_t) = \\
&= \sum_{i=-\infty}^{+\infty} Pr\{\alpha^2 \circ x_t = i\} Pr\{\alpha \circ \epsilon_{t+1} + \epsilon_{t+2} = x_{t+2} - i\} \\
&= \sum_{i=\max(0, x_{t+2} - 2m)}^{\min(x_{t+2}, x_t)} Pr\{\alpha^2 \circ x_t = i\} Pr\{\alpha \circ \epsilon_{t+1} + \epsilon_{t+2} = x_{t+2} - i\} \quad (2.10)
\end{aligned}$$

con $x_{t+2} \in \{0, 1, 2, \dots, (x_t + 2m)\}$. Nella sommatoria della (2.10) si ha che i

varia dal $\max(0, x_{t+2} - 2m)$ al $\min(x_{t+2}, x_t)$ considerando che $Pr\{\alpha^2 \circ x_t = i\}$ è non nulla per $i \in \{0, 1, 2, \dots, x_t\}$ e $Pr\{\alpha \circ \epsilon_{t+1} + \epsilon_{t+2} = x_{t+2} - i\}$ è non nulla per $x_{t+2} - i \in \{0, 1, 2, \dots, 2m\}$.

La distribuzione di $\alpha \circ \epsilon_{t+1} + \epsilon_{t+2}$ è a sua volta la convoluzione tra la distribuzione di $\alpha \circ \epsilon_{t+1}$ e ϵ_t , quindi il secondo termine della (2.10) risulta

$$\begin{aligned} Pr\{\alpha \circ \epsilon_{t+1} + \epsilon_{t+2} = x_{t+2} - i\} &= \\ &= \sum_{s=-\infty}^{+\infty} Pr\{\alpha \circ \epsilon_{t+1} = s\} Pr\{\epsilon_{t+2} = x_{t+2} - i - s\} = \\ &= \sum_{s=\max(0, x_{t+2}-i-m)}^{\min(m, x_{t+2}-i)} Pr\{\alpha \circ \epsilon_{t+1} = s\} Pr\{\epsilon_{t+2} = x_{t+2} - i - s\} \quad (2.11) \end{aligned}$$

nella sommatoria della (2.11) si ha che s varia dal $\max(0, x_{t+2} - i - m)$ al $\min(m, x_{t+2} - i)$ essendo $Pr\{\alpha \circ \epsilon_{t+1} = s\}$ non nulla per $s \in \{0, 1, 2, \dots, m\}$ e $Pr\{\epsilon_{t+2} = x_{t+2} - i - s\}$ non nulla per $x_{t+2} - i - s \in \{0, 1, 2, \dots, m\}$.

Inserendo la (2.11) nella (2.10) e sostituendo per ogni distribuzione di probabilità la relativa espressione si ottiene la distribuzione condizionale per $h = 2$

$$\begin{aligned} p(X_{t+2}|X_t) &= Pr(X_{t+2} = x_{t+2}|X_t = x_t) = \\ &= \sum_{i=\max(0, x_{t+2}-2m)}^{\min(x_{t+2}, x_t)} \binom{x_t}{i} \alpha^{2i} (1 - \alpha^2)^{x_t-i} \sum_{s=\max(0, x_{t+2}-i-m)}^{\min(m, x_{t+2}-i)} \binom{m}{s} \\ &\quad \times (\alpha p)^s (1 - \alpha p)^{m-s} \binom{m}{x_{t+2}-i-s} p^{x_{t+2}-i-s} (1 - p)^{m-x_{t+2}+i+s} \quad (2.12) \end{aligned}$$

con $x_{t+2} \in \{0, 1, 2, \dots, (x_t + 2m)\}$.

Le previsioni per $h = 1$ e $h = 2$ si ottengono sostituendo rispettivamente nella (2.9) e nella (2.12) ai parametri le loro stime e calcolando la mediana o la moda di queste distribuzioni. Le distribuzioni condizionali per valori di

h maggiori di 2 non sono state calcolate perchè le formulazioni risultavano molto pesanti; inoltre, queste non sono di particolare interesse dato che nei processi INAR(1), come nei classici AR(1), la correlazione decresce esponenzialmente con il ritardo temporale e quindi la previsione converge rapidamente alla marginale del processo. In ogni caso per ottenere previsioni per h maggiori di 2 è possibile approssimare la forma della distribuzione attraverso procedure di simulazione.

2.3 Innovazione Binomiale Negativa

2.3.1 Generalità

Un'altra possibile alternativa alla distribuzione di Poisson è la Binomiale Negativa.

Nei libri di testo esistono diverse formulazioni per la distribuzione Binomiale Negativa, nel seguito di questa tesi diremo che se ϵ_t si distribuisce come una Binomiale Negativa di parametri r e p , allora la sua distribuzione di probabilità è data da

$$Pr\{\epsilon_t = k\} = \frac{\Gamma(r+k)}{\Gamma(r)k!} p^r (1-p)^k \quad (2.13)$$

con $k \in \{1, 2, 3, \dots\}$.

Questa versione della Binomiale Negativa consente al parametro r di assumere valori continui e non soltanto discreti; ciò garantisce una maggiore flessibilità e, come si vedrà nel seguito, facilita la procedura di stima del modello da un punto di vista computazionale. La formulazione (2.13) ha lo svantaggio di perdere l'usuale interpretazione data dal conteggio del numero di insuccessi prima di raggiungere r successi, tuttavia questa limitazione non è importante per noi, visto il contesto di applicazione.

Rispetto alla situazione di equidispersione della Poisson, la Binomiale Negativa ha la caratteristica di permettere sovradisersione; ciò consente di considerare una differente casistica di forme della distribuzione del processo. Come nel caso Binomiale, la distribuzione marginale del processo non è nota, ma possono comunque essere sfruttati i risultati (1.3), (1.4) e (1.5) per calcolare i primi due momenti. La media marginale è pari a

$$\mu_x = E(X_t) = \frac{r(1-p)}{p(1-\alpha)} \quad (2.14)$$

la varianza marginale risulta

$$\sigma_x^2 = Var(X_t) = \frac{r(1-p)(\alpha p + 1)}{p^2(1-\alpha^2)} \quad (2.15)$$

mentre, la funzione di autocovarianza è

$$\gamma(k) = Cov(X_t, X_{t-k}) = \alpha^k \gamma(0) = \alpha^k \frac{r(1-p)(\alpha p + 1)}{p^2(1-\alpha^2)} \quad (2.16)$$

Dalla (2.14) e dalla (2.15) è facile vedere come, oltre all'innovazione, anche la distribuzione marginale del processo abbia sempre la varianza maggiore della media.

2.3.2 Procedura di stima

Con la specificazione Binomiale Negativa dell'innovazione si devono stimare tre parametri α , r e p . La funzione di verosimiglianza è data da

$$L(\alpha, r, p|x) = \left(\prod_{t=1}^n p(X_t|X_{t-1}) \right) p(X_0) \quad (2.17)$$

essendo X_t pari alla somma di due variabili indipendenti, la distribuzione di

$X_t|X_{t-1} = x_{t-1}$ è la convoluzione tra le distribuzioni di probabilità di $\alpha \circ x_{t-1}$ e ϵ_t

$$\begin{aligned}
p(X_t|X_{t-1}) &= Pr\{X_t = x_t|X_{t-1} = x_{t-1}\} = \\
&= Pr\{\alpha \circ X_{t-1} + \epsilon_t = x_t|X_{t-1} = x_{t-1}\} = \\
&= \sum_{i=-\infty}^{+\infty} Pr\{\alpha \circ x_{t-1} = i\}Pr\{\epsilon_t = x_t - i\} = \\
&= \sum_{i=0}^{\min(x_t, x_{t-1})} Pr\{\alpha \circ x_{t-1} = i\}Pr\{\epsilon_t = x_t - i\}
\end{aligned} \tag{2.18}$$

nella sommatoria della (2.18) si ha che i varia da zero al $\min(x_t, x_{t-1})$ considerando che $Pr\{\alpha \circ x_{t-1} = i\}$ è non nulla per $i \in \{0, 1, 2, \dots, x_{t-1}\}$ e $Pr\{\epsilon_t = x_t - i\}$ è non nulla per $x_t - i \in \{0, 1, 2, \dots\}$.

Sostituendo nella (2.18) per ogni distribuzione di probabilità la relativa espressione, si ottiene che

$$\begin{aligned}
p(X_t|X_{t-1}) &= \\
&= \sum_{i=0}^{\min(x_t, x_{t-1})} \binom{x_{t-1}}{i} \alpha^i (1 - \alpha)^{x_{t-1}-i} \left(\frac{\Gamma(r + x_t - i)}{\Gamma(r)(x_t - i)!} \right) p^r (1 - p)^{x_t - i} = \\
&= (1 - \alpha)^{x_{t-1}} p^r (1 - p)^{x_t} \sum_{i=0}^{\min(x_t, x_{t-1})} \binom{x_{t-1}}{i} \\
&\quad \times \frac{\Gamma(r + x_t - i)}{\Gamma(r)(x_t - i)!} \left(\frac{\alpha}{(1 - p)(1 - \alpha)} \right)^i
\end{aligned} \tag{2.19}$$

per $x_t \in \{0, 1, 2, \dots\}$.

Anche in questo caso la funzione di probabilità marginale del processo non è nota e la distribuzione di X_0 nella (2.17) non può essere ricavata in maniera esplicita. Si utilizza quindi la funzione di verosimiglianza condizionata a

$X_0 = x_0$ che è: $L(\alpha, r, p|x_0) = \prod_{t=1}^n p(X_t|X_{t-1})$, da cui si ricava la funzione di log-verosimiglianza condizionata

$$\begin{aligned}
l(\alpha, r, p|x_0) &= \log(L(\alpha, r, p|x_0)) = \sum_{t=1}^n \log(p(X_t|X_{t-1})) = \\
&= \left(\sum_{t=1}^n x_{t-1} \right) \log(1 - \alpha) + nr \log(p) + \left(\sum_{t=1}^n x_t \right) \log(1 - p) + \\
&+ \sum_{t=1}^n \log \left(\sum_{i=0}^{\min(x_t, x_{t-1})} \binom{x_{t-1}}{i} \frac{\Gamma(r + x_t - i)}{\Gamma(r)(x_t - i)!} \left(\frac{\alpha}{(1-p)(1-\alpha)} \right)^i \right)
\end{aligned} \tag{2.20}$$

Le stime di massima verosimiglianza sono i valori di α , r e p che massimizzano la (2.20). Non essendo ricavabile una soluzione esplicita per questo problema di massimo, si deve procedere per via numerica attraverso un algoritmo di ottimizzazione che massimizzi la log-verosimiglianza tenendo anche conto dei valori che possono essere assunti dai parametri, infatti, $\alpha \in [0, 1)$, $r \in R^+$ e $p \in (0, 1)$.

2.3.3 Previsioni

Come nel caso Binomiale non è possibile ricavare una forma generica della funzione di probabilità h passi in avanti dato che la distribuzione di $\sum_{i=0}^{h-1} \alpha^i \circ \epsilon_{t+h-i}$ nella (1.10) non ha una formulazione generale al variare di h . In particolare, calcolando la funzione generatrice dei momenti in modo analogo a (1.11) e (2.8), è possibile far vedere come $\alpha^i \circ \epsilon_{t+h-i}$ non abbia una distribuzione nota. Si può comunque procedere calcolando la funzione di probabilità per $h = 1$ e $h = 2$.

La distribuzione per $h = 1$ di $X_{t+h}|X_t = x_t$ è facilmente ricavabile sfruttando il risultato (2.19)

$$\begin{aligned}
p(X_{t+1}|X_t) &= \\
&= (1-\alpha)^{x_t} p^r (1-p)^{x_{t+1}} \sum_{i=0}^{\min(x_{t+1}, x_t)} \binom{x_t}{i} \\
&\quad \times \left(\frac{\Gamma(r+x_{t+1}-i)}{\Gamma(r)(x_{t+1}-i)!} \right) \left(\frac{\alpha}{(1-p)(1-\alpha)} \right)^i \quad (2.21)
\end{aligned}$$

per $x_{t+1} \in \{0, 1, 2, \dots\}$.

Per $h = 2$, in riferimento alla (1.10), la $p(X_{t+h}|X_t = x_t)$ è la distribuzione della somma di tre variabili aleatorie indipendenti: $\alpha^2 \circ x_t$ che è una Binomiale di parametri x_t e α^2 ; $\alpha \circ \epsilon_{t+1}$ che ha una distribuzione non nota; e ϵ_{t+2} che è una Binomiale Negativa di parametri r e p .

La $p(X_{t+2}|X_t)$ è pari alla convoluzione tra la distribuzione di $\alpha^2 \circ x_t$ e la distribuzione di $\alpha \circ \epsilon_{t+1} + \epsilon_t$

$$\begin{aligned}
p(X_{t+2}|X_t) &= Pr(X_{t+2} = x_{t+2}|X_t = x_t) = \\
&= \sum_{i=-\infty}^{+\infty} Pr\{\alpha^2 \circ x_t = i\} Pr\{\alpha \circ \epsilon_{t+1} + \epsilon_{t+2} = x_{t+2} - i\} \\
&= \sum_{i=0}^{\min(x_{t+2}, x_t)} Pr\{\alpha^2 \circ x_t = i\} Pr\{\alpha \circ \epsilon_{t+1} + \epsilon_{t+2} = x_{t+2} - i\} \quad (2.22)
\end{aligned}$$

con $x_{t+2} \in \{0, 1, 2, \dots\}$. Nella sommatoria della (2.22) si ha che i varia da zero al $\min(x_{t+2}, x_t)$ considerando che $Pr\{\alpha^2 \circ x_t = i\}$ è non nulla per $i \in \{0, 1, 2, \dots, x_t\}$ e $Pr\{\alpha \circ \epsilon_{t+1} + \epsilon_{t+2} = x_{t+2} - i\}$ è non nulla per $x_{t+2} - i \in \{0, 1, 2, \dots\}$.

La distribuzione di $\alpha \circ \epsilon_{t+1} + \epsilon_{t+2}$ è a sua volta la convoluzione tra la distribuzione di $\alpha \circ \epsilon_{t+1}$ e ϵ_t , pertanto, il secondo termine della (2.22) risulta

$$\begin{aligned}
& Pr\{\alpha \circ \epsilon_{t+1} + \epsilon_{t+2} = x_{t+2} - i\} = \\
& = \sum_{j=-\infty}^{+\infty} Pr\{\alpha \circ \epsilon_{t+1} = j\} Pr\{\epsilon_{t+2} = x_{t+2} - i - j\} = \\
& = \sum_{j=0}^{x_{t+2}-i} Pr\{\alpha \circ \epsilon_{t+1} = j\} Pr\{\epsilon_{t+2} = x_{t+2} - i - j\} \quad (2.23)
\end{aligned}$$

nella sommatoria della (2.23) si ha che j varia da zero a $x_{t+2} - i$ considerando che $Pr\{\alpha \circ \epsilon_{t+1} = j\}$ è non nulla per $j \in \{0, 1, 2, \dots\}$ e $Pr\{\epsilon_{t+2} = x_{t+2} - i - j\}$ è non nulla per $x_{t+2} - i - j \in \{0, 1, 2, \dots\}$.

La distribuzione di probabilità di $\alpha \circ \epsilon_{t+1}$ non è nota, ma può essere ricavata dato che è una Binomiale con parametro di conteggio aleatorio e distribuito come una Binomiale negativa. La $Pr\{\alpha \circ \epsilon_{t+1} = j\}$ nella (2.23) è quindi pari a

$$\begin{aligned}
& Pr\{\alpha \circ \epsilon_{t+1} = j\} = \\
& = \sum_{s=-\infty}^{+\infty} Pr\{\alpha \circ \epsilon_{t+1} = j | \epsilon_{t+1} = s\} Pr\{\epsilon_{t+1} = s\} = \\
& = \sum_{s=j}^{+\infty} Pr\{\alpha \circ s = j\} Pr\{\epsilon_{t+1} = s\} \quad (2.24)
\end{aligned}$$

nella sommatoria della (2.24) si ha che s varia da j a $+\infty$ considerando che $Pr\{\alpha \circ s = j\}$ è non nulla per $s \in \{j, j+1, j+2, \dots\}$ e $Pr\{\epsilon_{t+1} = s\}$ è non nulla per $s \in \{0, 1, 2, \dots\}$.

Inserendo la (2.24) nella (2.23) e la (2.23) nella (2.22), e sostituendo per ogni distribuzione di probabilità la relativa espressione si ottiene la distribuzione condizionale per $h = 2$

$$\begin{aligned}
p(X_{t+2}|X_t) &= Pr(X_{t+2} = x_{t+2} | X_t = x_t) = \\
&= \sum_{i=0}^{\min(x_{t+2}, x_t)} \binom{x_t}{i} \alpha^{2i} (1 - \alpha^2)^{x_t - i}
\end{aligned}$$

$$\begin{aligned}
& \times \sum_{j=0}^{x_{t+2}-i} \frac{\Gamma(r+x_{t+2}-i-j)}{\Gamma(r)(x_{t+2}-i-j)!} p^r (1-p)^{x_{t+2}-i-j} \\
& \times \sum_{s=j}^{+\infty} \binom{s}{j} \alpha^j (1-\alpha)^{s-j} \frac{\Gamma(r+s)}{\Gamma(r)s!} p^r (1-p)^s \quad (2.25)
\end{aligned}$$

con $x_{t+2} \in \{0, 1, 2, \dots\}$.

Le previsioni puntuali per $h = 1$ e $h = 2$ si ottengono sostituendo rispettivamente nella (2.21) e nella (2.25) le relative stime dei parametri e calcolando la mediana o la moda di queste distribuzioni. In generale, si possono utilizzare le espressioni (2.21) e (2.25) per ottenere la previsione della probabilità di un evento di interesse. Le distribuzioni condizionali per valori di h maggiori di 2 non sono state calcolate per gli stessi motivi già illustrati nel paragrafo relativo all'innovazione Binomiale.

Capitolo 3

Esperimento Monte Carlo

3.1 Informazioni generali

In questo capitolo viene condotto un esperimento di simulazione Monte Carlo per testare le procedure di stima e previsione presentate in precedenza.

L'esperimento è suddiviso in due parti. La prima parte tratta gli stimatori di massima verosimiglianza e ha lo scopo di ottenere informazioni riguardanti la distorsione e la variabilità degli stimatori sia al variare della numerosità campionaria sia al variare dei valori assunti dai parametri. La seconda parte è un confronto tra i diversi modelli trattati in termini di qualità delle previsioni e ha, principalmente, lo scopo di verificare se l'utilizzo dei modelli INAR(1) con specificazioni diverse dalla Poisson possa essere utile in fase previsiva.

Lo studio di simulazione è stato condotto implementando col *software* statistico R le funzioni di log-verosimiglianza e le distribuzioni condizionali per le previsioni, su serie generate da modelli INAR(1) con diversi valori dei parametri e diverse specificazioni del termine d'errore. Le stime di massima verosimiglianza dei parametri sono state ottenute utilizzando un ottimizzatore numerico che implementa l'algoritmo di Nelder e Mead (1965).

3.2 Stimatori di massima verosimiglianza

3.2.1 Motivazioni e obbiettivi

Come già introdotto in precedenza, questa sezione ha lo scopo di valutare da un punto di vista qualitativo gli stimatori di massima verosimiglianza sia al variare della numerosità campionaria sia al variare dei valori assunti dai parametri. L'interesse è focalizzato sulle usuali caratteristiche degli stimatori e cioè: consistenza e distorsione in campioni finiti e l'eventuale velocità di convergenza alla distribuzione Normale. Si fa, inoltre, presente che questo esperimento ha un fine puramente illustrativo e non mira al conseguimento di particolari conclusioni dato che l'unica procedura utilizzata è la massima verosimiglianza e non è stato fatto un confronto tra diverse tecniche di stima. Per questo motivo i risultati delle simulazioni sono presentati soltanto graficamente e le considerazioni a riguardo sono di carattere generale. Nonostante ciò si ritiene comunque utile esporre questi risultati dato che in letteratura non è presente alcuna applicazione sulla stima di massima verosimiglianza dei modelli INAR(1) con il termine d'errore Binomiale e Binomiale Negativa. Nel seguito, sono esposti i risultati delle simulazioni suddivise per le tre differenti specificazioni dell'errore del processo trattate in questa tesi.

3.2.2 Innovazione Poisson

In questo primo caso l'innovazione del processo è Poisson e i parametri da stimare con il metodo della massima verosimiglianza sono α e λ .

Dalla figura 3.2 si può vedere come il valore atteso degli stimatori converga al vero valore del parametro e la varianza converga a zero all'aumentare della numerosità campionaria. In campioni finiti sembra essere presente una lieve distorsione; in particolare, sembra che gli stimatori di massima verosimiglianza mediamente sottostimino il parametro α e sovrastimino il parametro λ . Questo è messo in evidenza sia in figura 3.2 che in figura 3.3 e 3.4, dove si vede come, anche al variare dei parametri, la distorsione sia sistematicamente positiva per lo stimatore di λ e negativa per lo stimatore di α .

Osservando la figura 3.1 si nota come la distribuzione degli stimatori, in accordo con le proprietà asintotiche, converga rapidamente alla Normale e che, anche per numerosità campionarie relativamente basse, l'approssimazione Gaussiana sia da considerarsi accettabile.

Dalla figura 3.3 si vede come l'errore quadratico medio (MSE) dello stimatore del parametro α dipenda fortemente dal valore che esso assume e, in particolare, per valori bassi di questo l'MSE sia più elevato. Allo stesso modo, anche la precisione dello stimatore del parametro λ sembra dipendere dal valore che questo assume, con l'MSE che aumenta per valori elevati (figura 3.4).

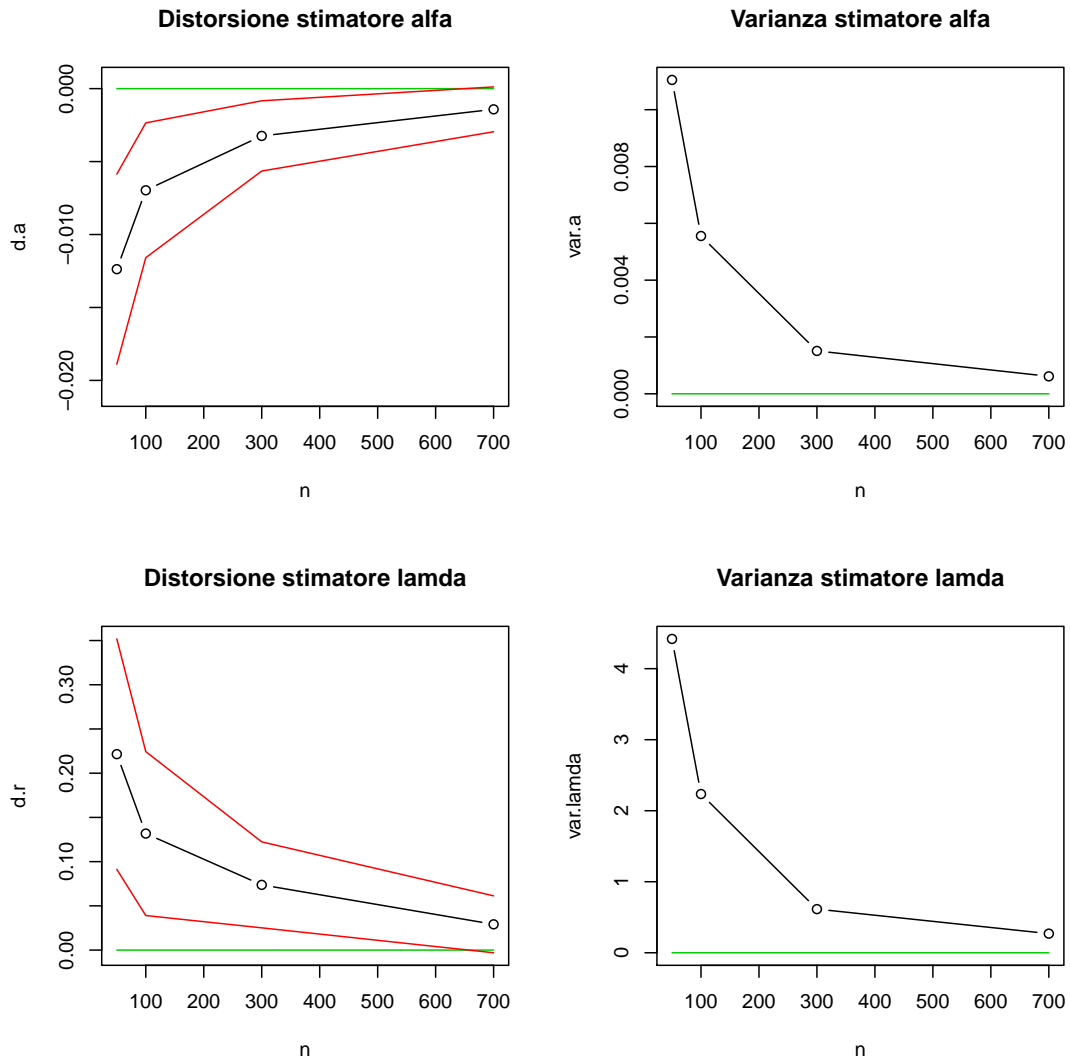


Figura 3.1: Nei grafici di sx: distorsione degli stimatori al variare della numerosità campionaria con intervalli di confidenza al 95%. Nei grafici di dx: varianza degli stimatori al variare della numerosità campionaria. 1000 simulazioni con: $(\alpha = 0.5, \lambda = 10)$.

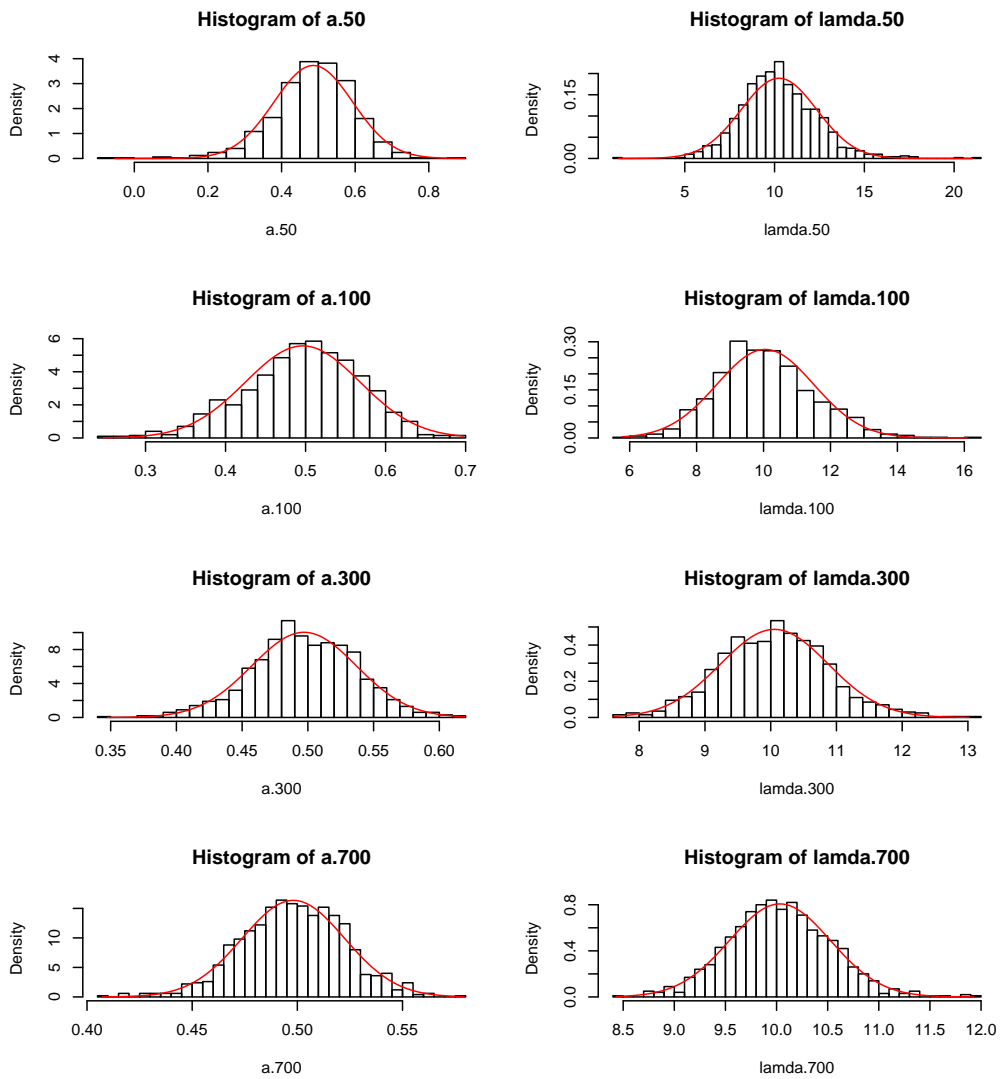


Figura 3.2: Istogrammi degli stimatori dei parametri ($\alpha = 0.5, \lambda = 10$) basati su 1000 simulazioni per numerosità 50, 100, 300 e 700. In rosso la densità della distribuzione Normale.

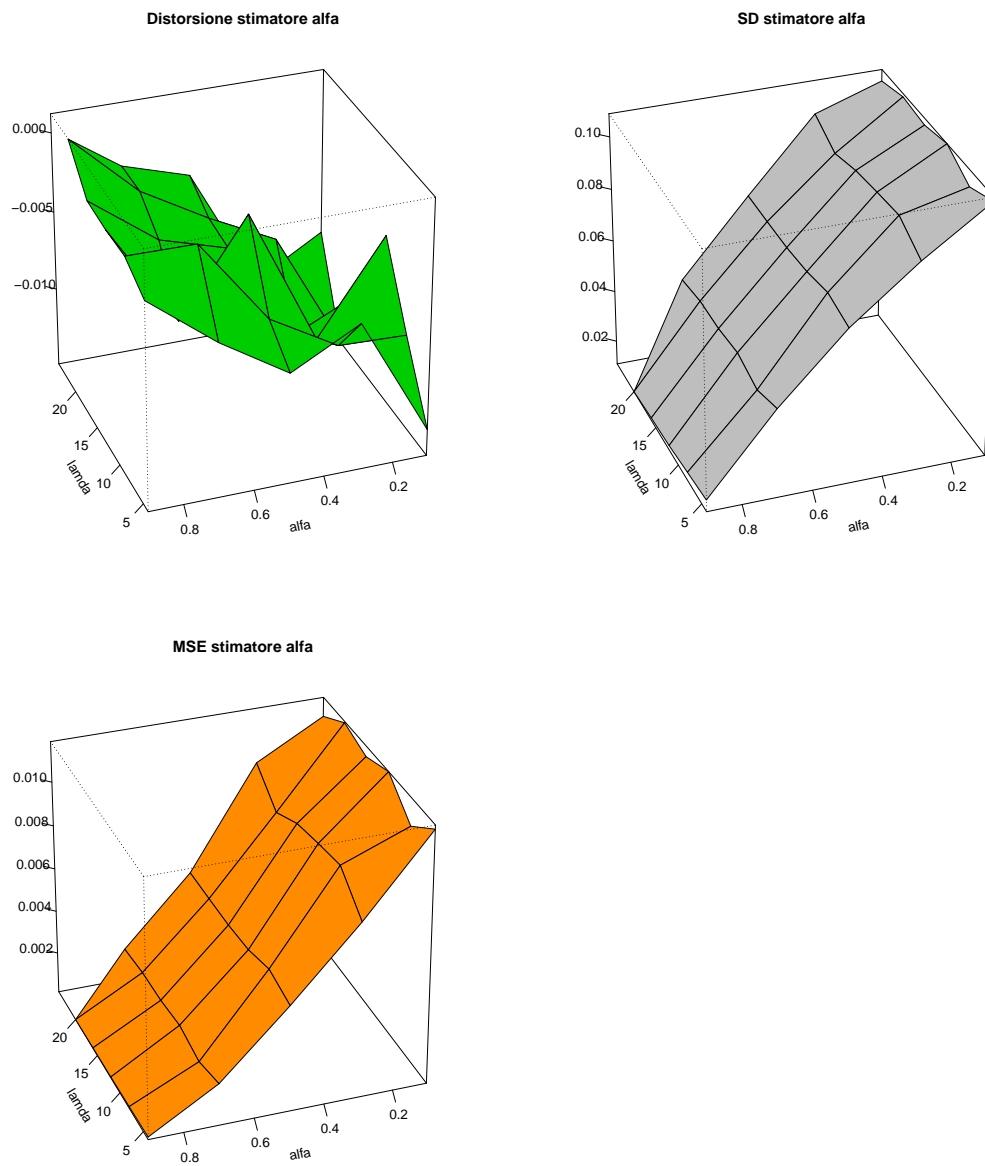


Figura 3.3: Nel grafico in alto a sx: distorsione dello stimatore di α al variare di α e λ . Nel grafico in alto a dx: deviazione standard dello stimatore di α al variare di α e λ . Nel grafico in basso a sx: MSE dello stimatore di α al variare di α e λ . 1000 simulazioni con numerosità 100.

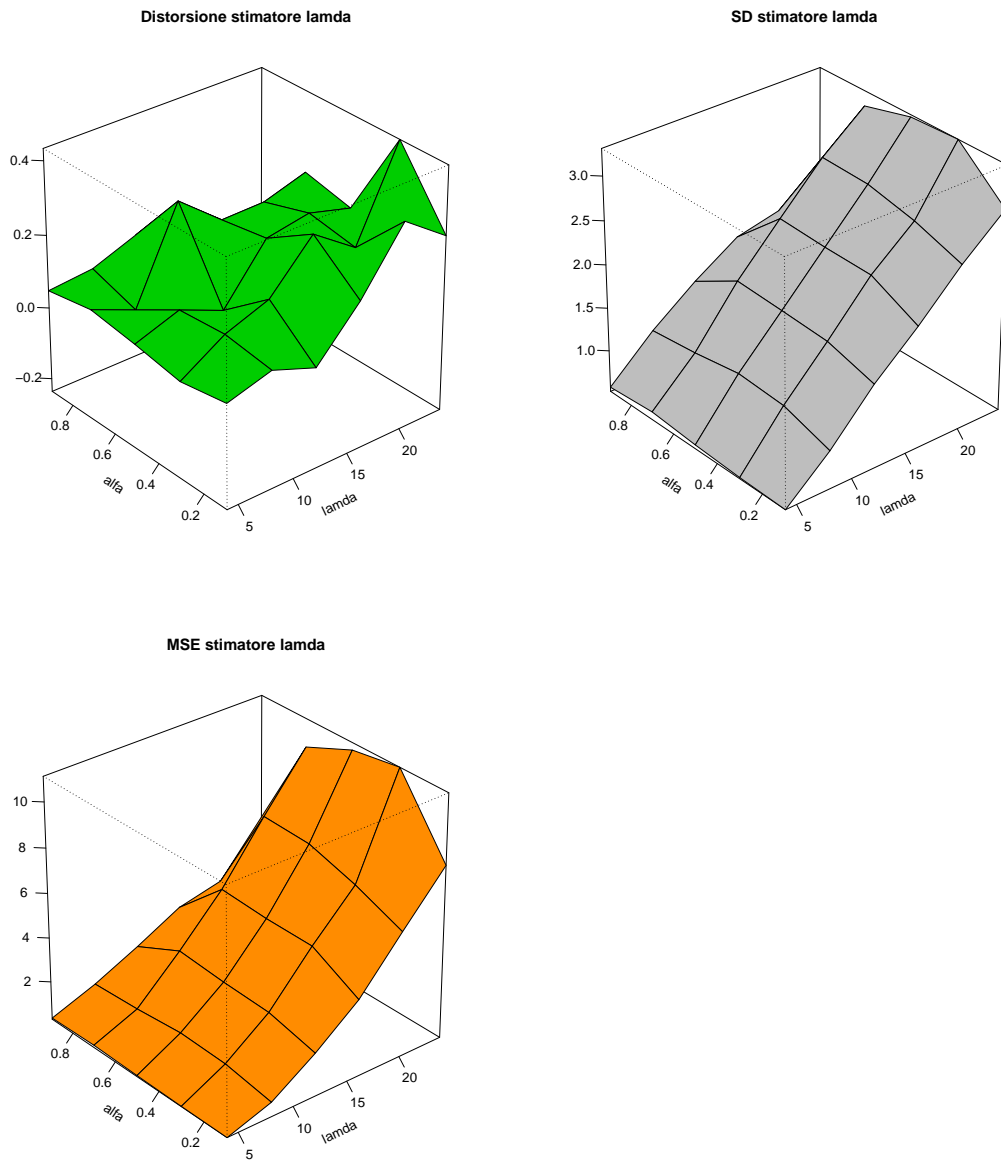


Figura 3.4: Nel grafico in alto a sx: distorsione dello stimatore di λ al variare di α e λ . Nel grafico in alto a dx: deviazione standard dello stimatore di λ al variare di α e λ . Nel grafico in basso a sx: MSE dello stimatore di λ al variare di α e λ . 1000 simulazioni con numerosità 100.

3.2.3 Innovazione Binomiale

Quando il termine d'errore ϵ_t , nella (1.1) è distribuito come una Binomiale con entrambi i parametri ignoti, i parametri da stimare con il metodo della massima verosimiglianza sono α , m e p . Il parametro m è discreto e inoltre, come già detto nel secondo capitolo, le usuali proprietà asintotiche degli stimatori di massima verosimiglianza non sono garantite.

La figura 3.5 evidenzia come gli stimatori dei tre parametri sembrano essere consistenti dato che la varianza converge a zero e il valore atteso converge al vero valore del parametro all'aumentare della numerosità campionaria.

In campioni finiti si nota la presenza di un'evidente distorsione, con il parametro α mediamente sottostimato e i parametri m e p mediamente sovrastimati. Il segno della distorsione sembra costante al variare dei valori assunti dai parametri, mentre il grado di distorsione sembra da essi influenzato (figura 3.7, 3.8 e 3.9).

Osservando la figura 3.9 si vede come lo stimatore di p non sia ben approssimabile dalla distribuzione Gaussiana nemmeno a elevate numerosità campionarie (si veda istogramma a numerosità 700), mentre quello di α sia approssimabile in maniera migliore. Si evidenzia inoltre la distribuzione discreta dello stimatore di m .

Le figure 3.7, 3.8 e 3.9 mostrano come varia la precisione degli stimatori dei tre parametri al variare di α (parametro che regola la dipendenza tra le variabili del processo) e di p (parametro che regola il grado di sottodispersione nella distribuzione Binomiale).

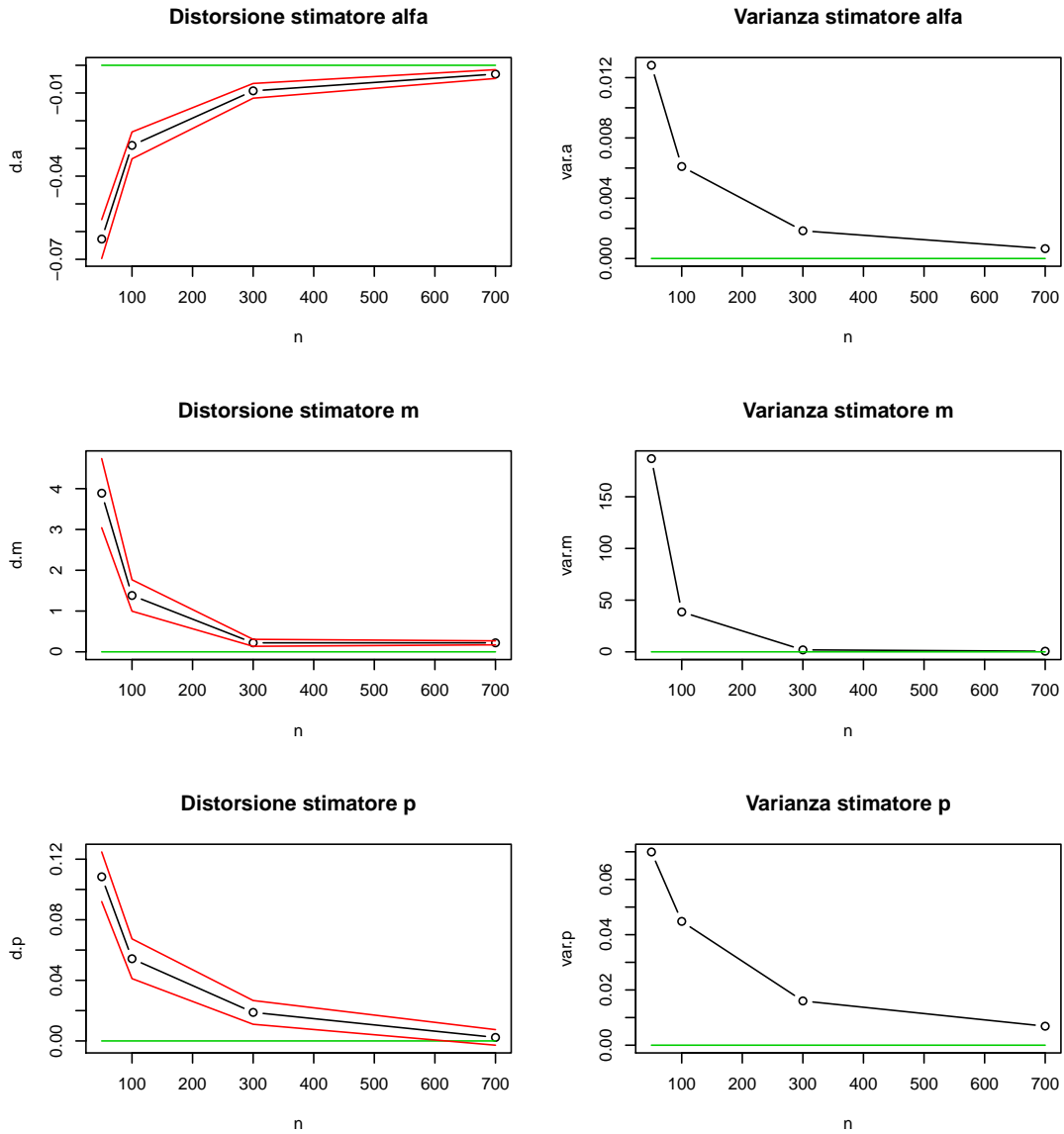


Figura 3.5: Nei grafici di sx: distorsione degli stimatori al variare della numerosità campionaria con intervalli di confidenza al 95%. Nei grafici di dx: varianza degli stimatori al variare della numerosità campionaria. 1000 simulazioni con: $(\alpha = 0.7, m = 5, p = 0.5)$.

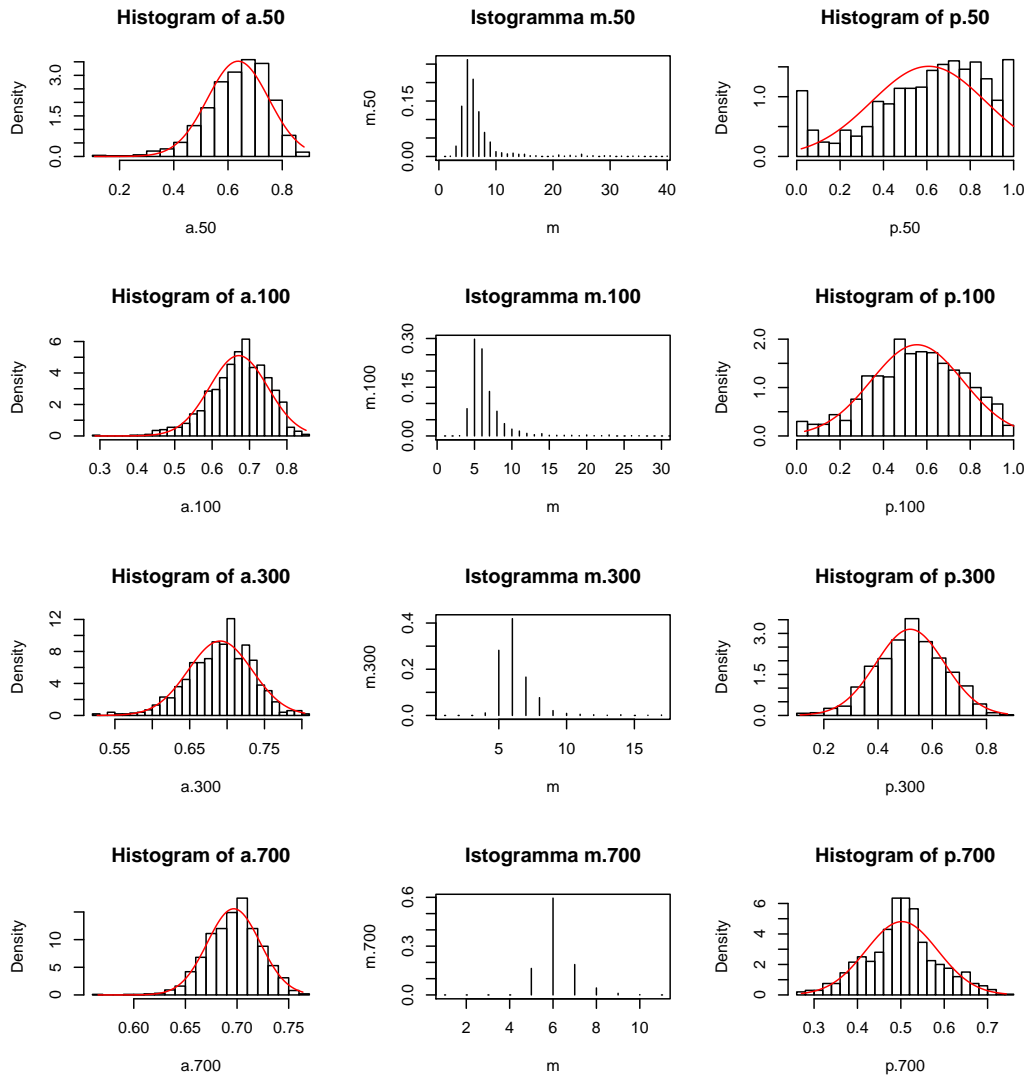


Figura 3.6: Istogrammi degli stimatori dei parametri ($\alpha = 0.7, m = 6, p = 0.5$) basati su 1000 simulazioni per numerosità 50, 100, 300 e 700. In rosso la densità della distribuzione Normale.

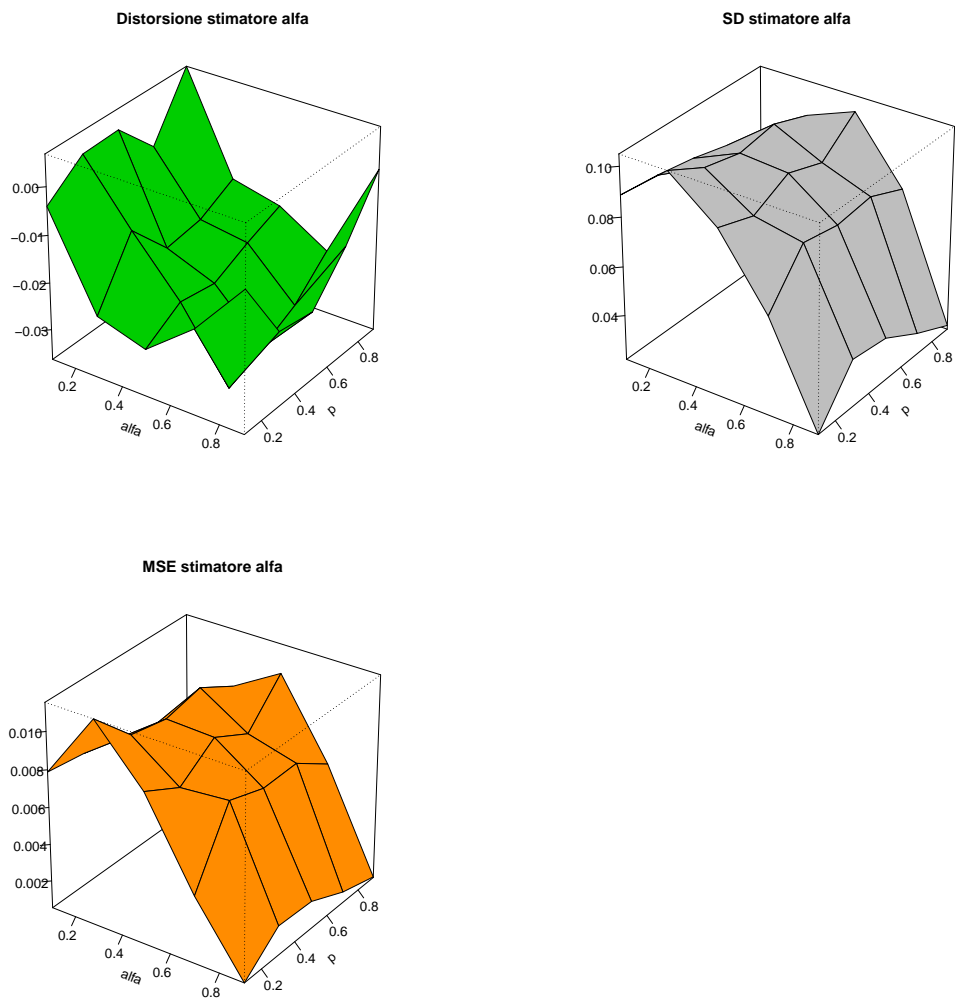


Figura 3.7: Nel grafico in alto a sx: distorsione dello stimatore di α al variare di α e p . Nel grafico in alto a dx: deviazione standard dello stimatore di α al variare di α e p . Nel grafico in basso a sx: MSE dello stimatore di α al variare di α e p . 500 simulazioni con numerosità 100 e $m = 6$.

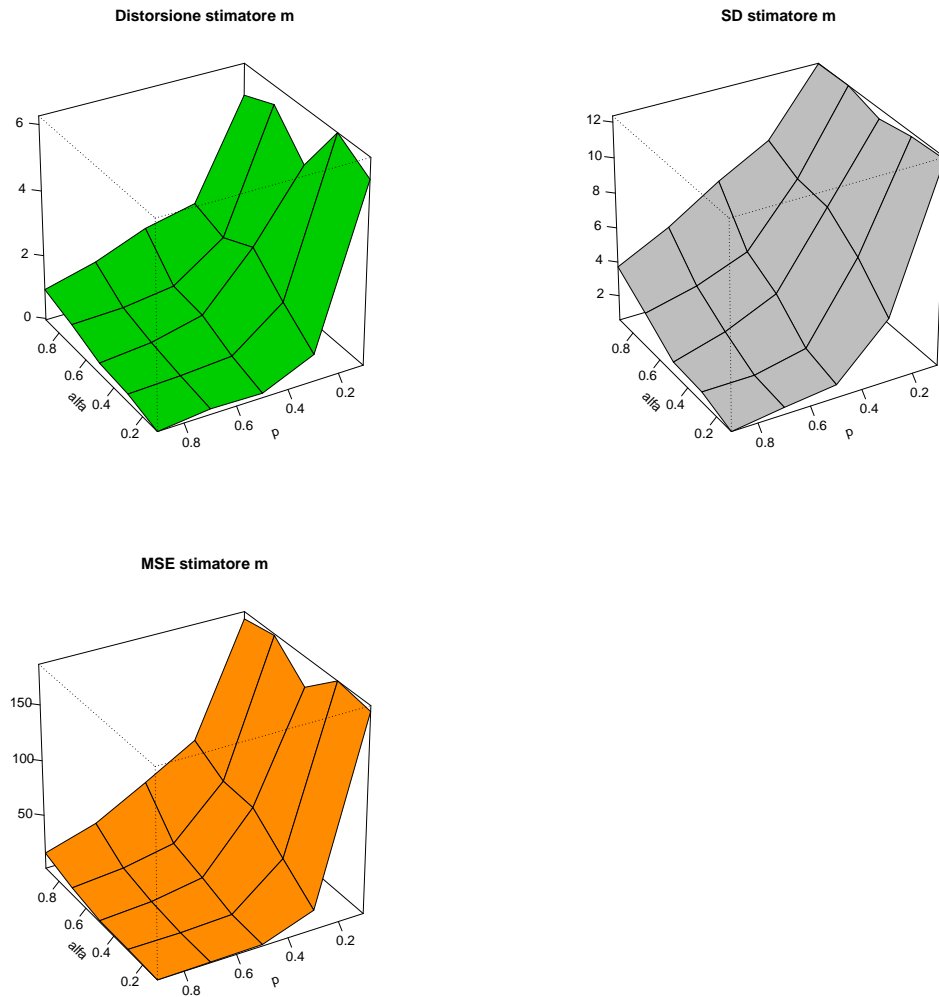


Figura 3.8: Nel grafico in alto a sx: distorsione dello stimatore di m al variare di α e p . Nel grafico in alto a dx: deviazione standard dello stimatore di m al variare di α e p . Nel grafico in basso a sx: MSE dello stimatore di m al variare di α e p . 500 simulazioni con numerosità 100 e $m = 6$.

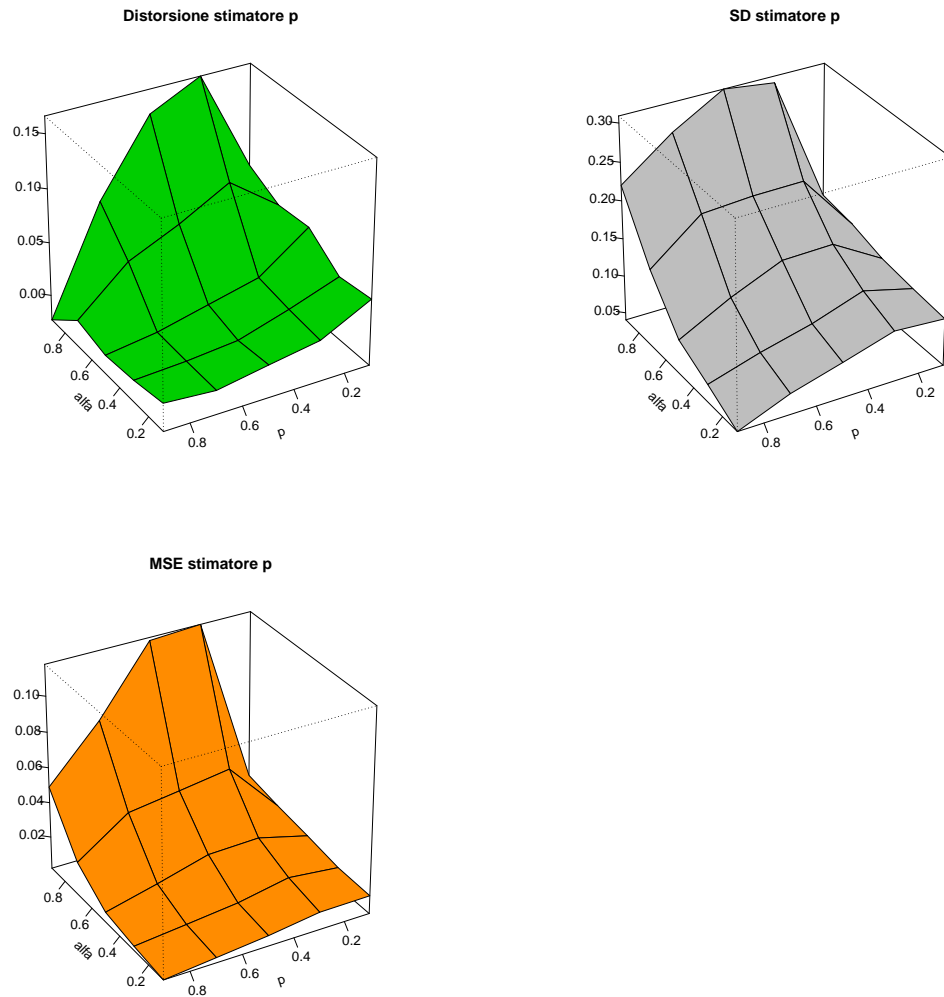


Figura 3.9: Nel grafico in alto a sx: distorsione dello stimatore di p al variare di α e p . Nel grafico in alto a dx: deviazione standard dello stimatore di p al variare di α e p . Nel grafico in basso a sx: MSE dello stimatore di p al variare di α e p . 500 simulazioni con numerosità 100 e $m = 6$.

3.2.4 Innovazione Binomiale Negativa

In questo ultimo caso, l'innovazione del processo è distribuita come una Binomiale Negativa e i parametri da stimare con il metodo della massima verosimiglianza sono α , r e p .

Dalla figura 3.10, si evidenzia come gli stimatori dei tre parametri sembrano essere consistenti, infatti, la varianza converge a zero e il valore atteso converge al vero valore del parametro all'aumentare della numerosità campionaria.

In campioni finiti si nota un evidente distorsione, con il parametro α mediamente sottostimato e i parametri r e p mediamente sovrastimati. Anche in questo caso, il segno della distorsione non sembra dipendere dai valori assunti dai parametri, mentre il grado di distorsione pare da questi influenzato (figura 3.12, 3.13 e 3.14).

Osservando la figura 3.11, si vede come la distribuzione degli stimatori dei tre parametri sembra avere diverse velocità di convergenza alla normalità. Lo stimatore di α può essere considerato approssimabile ad una normale già a basse numerosità campionarie, mentre questo non vale per gli stimatori di r e p .

Infine, le figure 3.12, 3.13 e 3.14 mostrano come varia la precisione dei tre stimatori al variare di α (parametro che regola la dipendenza tra le variabili del processo) e di p (parametro che regola la sovradisersione della distribuzione Binomiale Negativa).

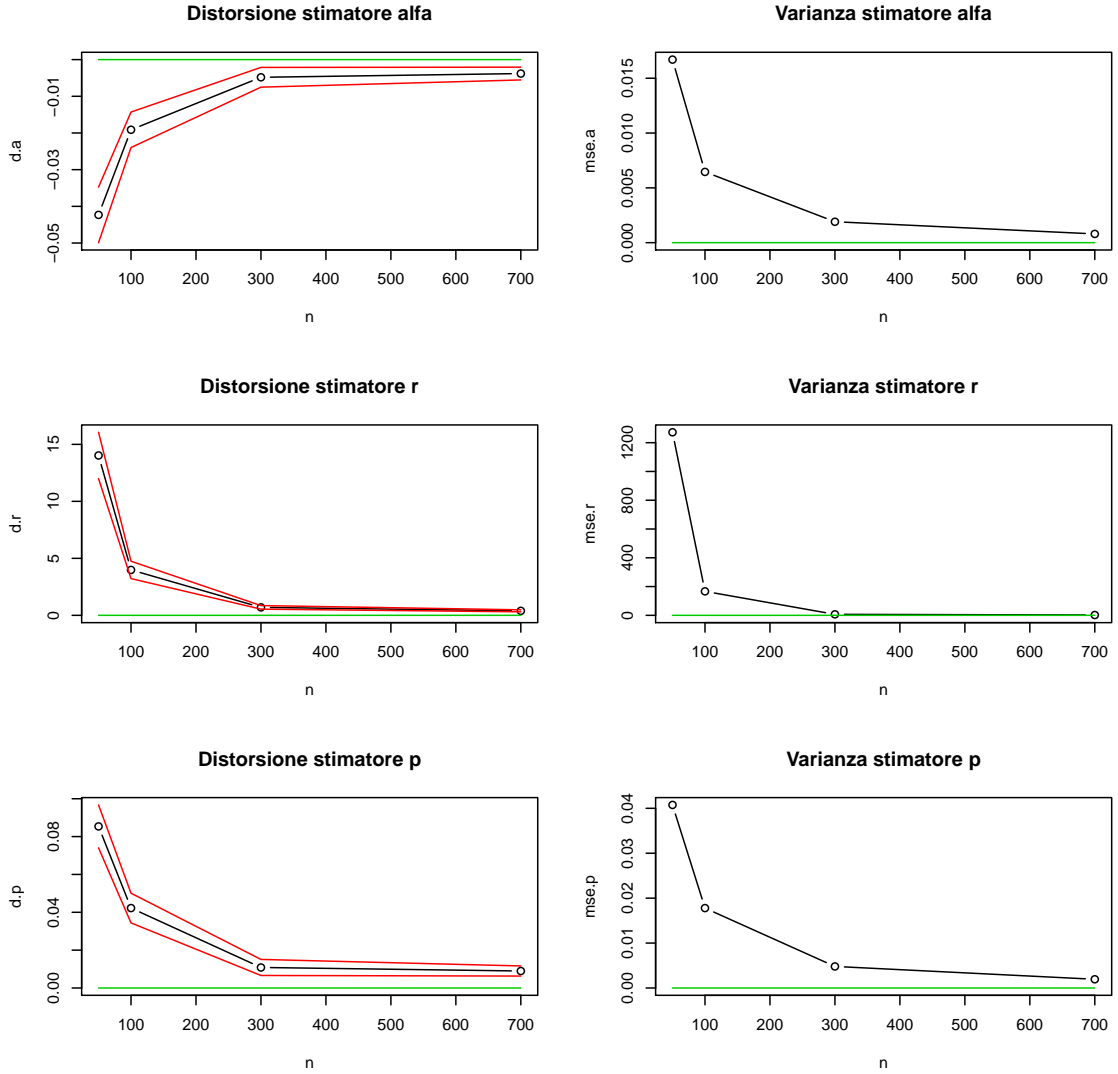


Figura 3.10: Nei grafici di sx: distorsione degli stimatori al variare della numerosità campionaria con intervalli di confidenza al 95%. Nei grafici di dx: varianza degli stimatori al variare della numerosità campionaria. 1000 simulazioni con: $(\alpha = 0.5, r = 6, p = 0.5)$.

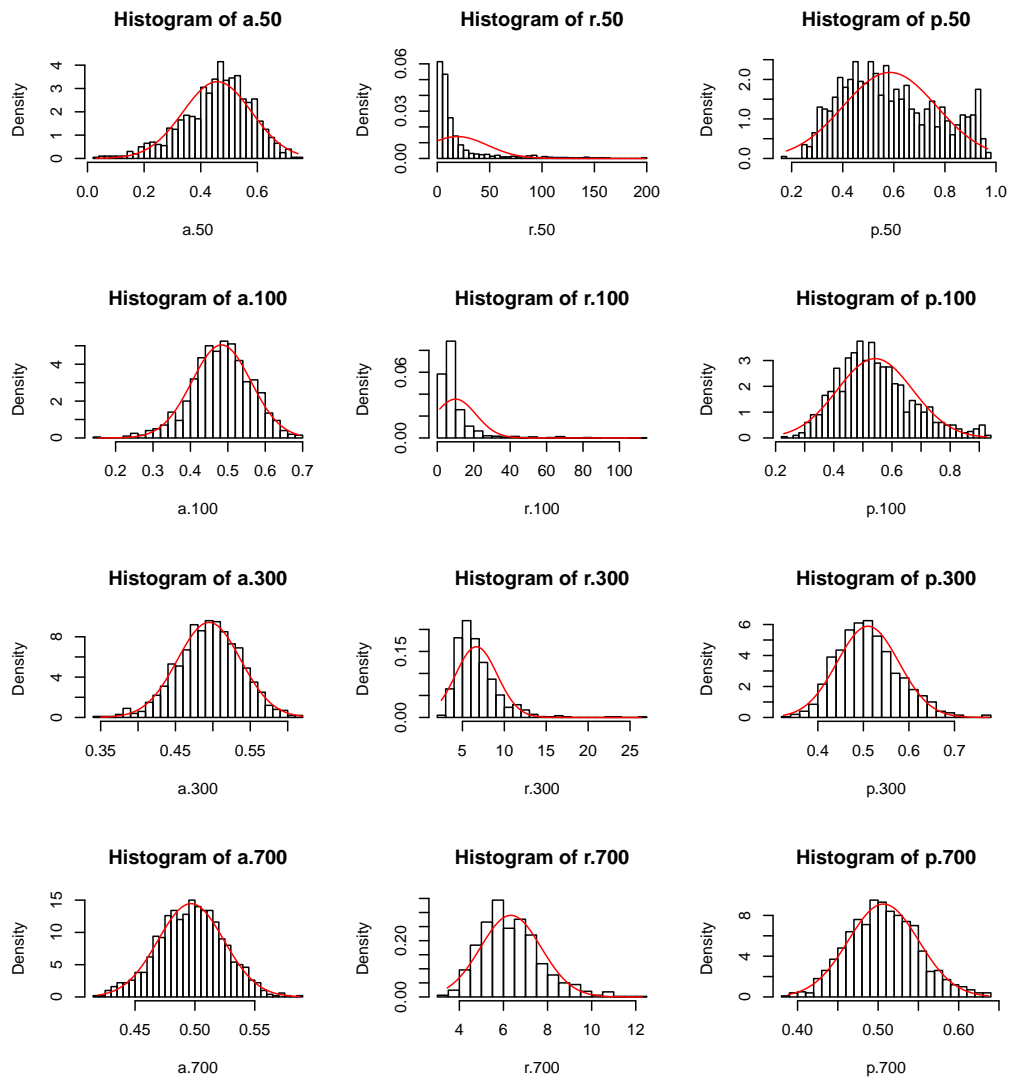


Figura 3.11: Istogrammi degli stimatori dei parametri ($\alpha = 0.5, r = 6, p = 0.5$) basati su 1000 simulazioni per numerosità 50, 100, 300 e 700. In rosso la densità della distribuzione Normale

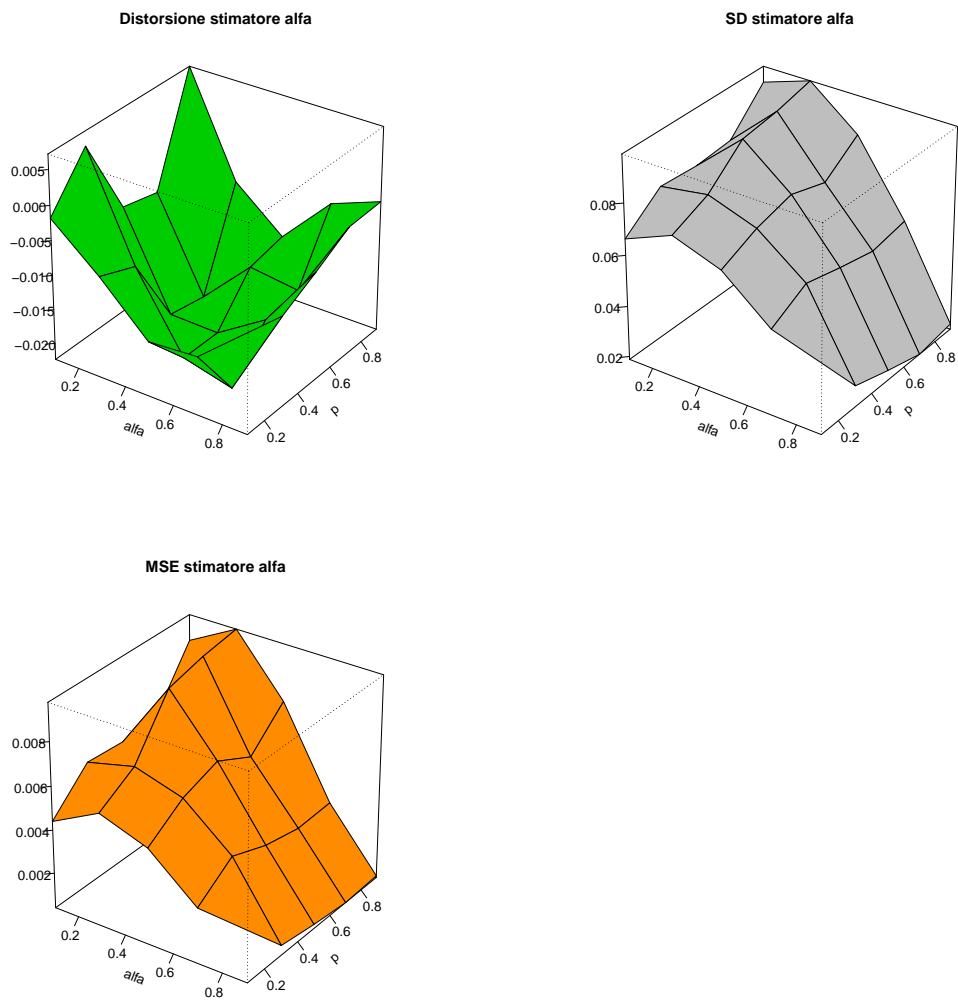


Figura 3.12: Nel grafico in alto a sx: distorsione dello stimatore di α al variare di α e p . Nel grafico in alto a dx: deviazione standard dello stimatore di α al variare di α e p . Nel grafico in basso a sx: MSE dello stimatore di α al variare di α e p . 500 simulazioni con numerosità 100 e $r = 4$.

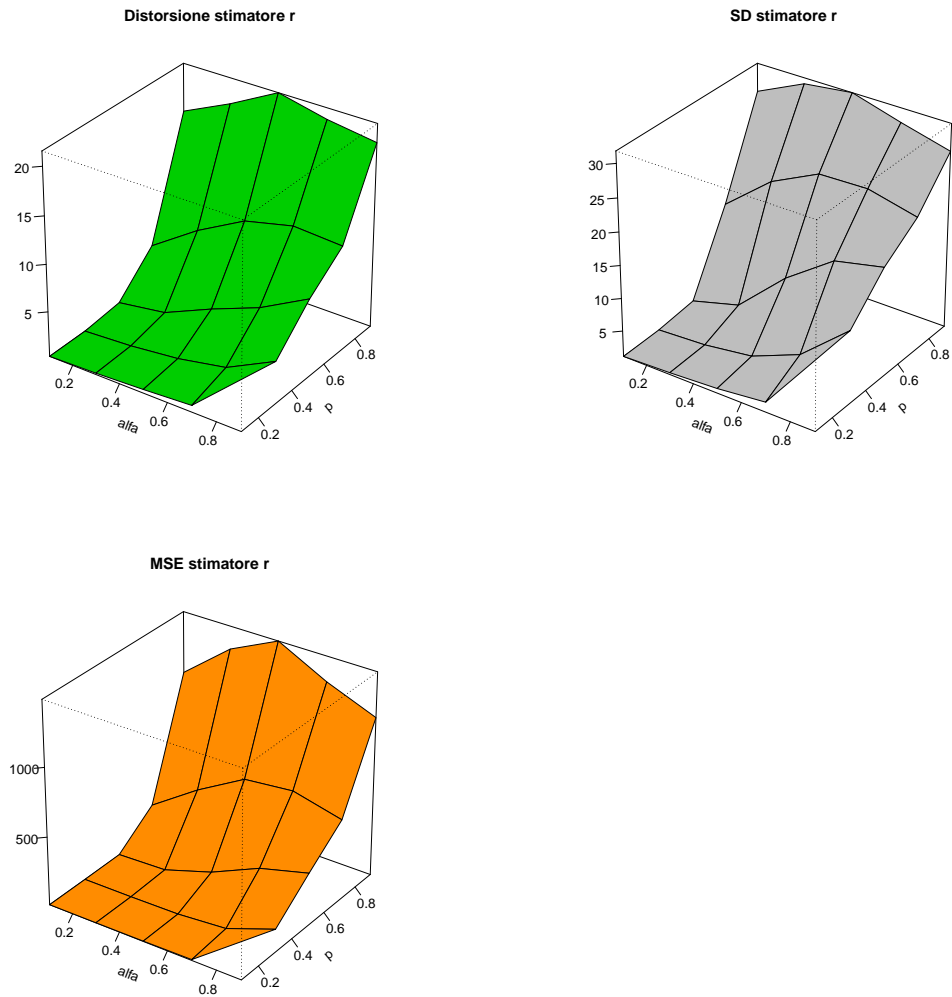


Figura 3.13: Nel grafico in alto a sx: distorsione dello stimatore di r al variare di α e p . Nel grafico in alto a dx: deviazione standard dello stimatore di r al variare di α e p . Nel grafico in basso a sx: MSE dello stimatore di r al variare di α e p . 500 simulazioni con numerosità 100 e $r = 4$.

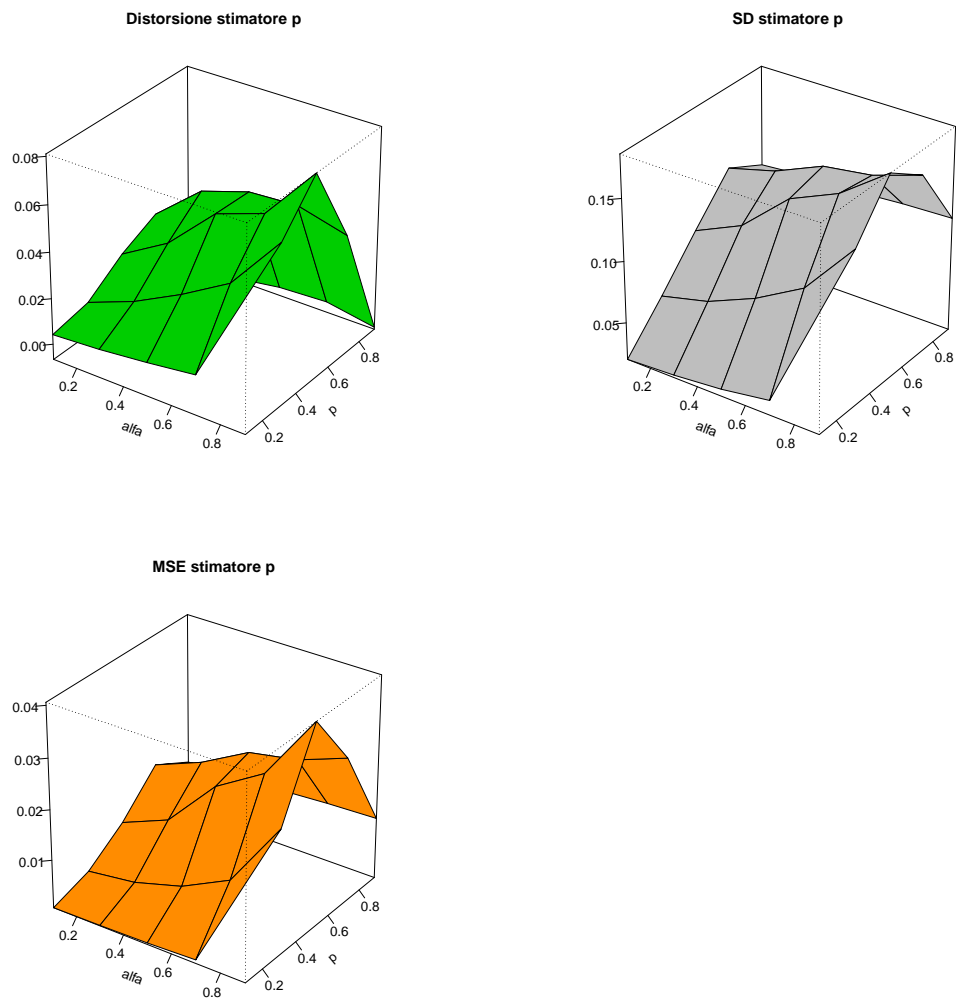


Figura 3.14: Nel grafico in alto a sx: distorsione dello stimatore di p al variare di α e p . Nel grafico in alto a dx: devizione standard dello stimatore di p al variare di α e p . Nel grafico in basso a sx: MSE dello stimatore di p al variare di α e p . 500 simulazioni con numerosità 100 e $r = 4$.

3.3 Previsioni

3.3.1 Obbiettivi e criteri utilizzati

L'obiettivo principale di questa tesi è quello di ottenere delle alternative al modello PoINAR(1) specificando il termine d'errore del processo con distribuzioni diverse dalla Poisson. Le distribuzioni scelte sono la distribuzione Binomiale e la distribuzione Binomiale Negativa, con la prima che implica sottodispersione (media maggiore della varianza) e la seconda sovradispersione (media minore della varianza). Chiaramente, affinché queste alternative possano essere considerate valide, dovrebbero, almeno in alcune circostanze, garantire delle *performance* migliori rispetto al modello Poisson. In questa sezione lo studio di simulazione condotto ha proprio lo scopo di capire se l'utilizzo di questi modelli alternativi possa essere utile, portando a delle migliori *performance* in fase previsiva. L'esperimento si basa sul confronto tra dei modelli correttamente specificati e dei modelli mal specificati, verificando così se ci siano sensibili differenze in termini di qualità delle previsioni.

Lo studio è suddiviso in tre parti: nella prima parte il confronto è fatto tra il modello PoINAR(1) correttamente specificato e il modello AR(1) mal specificato; nella seconda parte il confronto è tra il modello INAR(1) con errore Binomiale e i modelli PoINAR(1) e AR(1) mal specificati; e, nella terza parte, il confronto è fatto tra il modello INAR(1) con errore Binomiale Negativa correttamente specificato e i modelli PoINAR(1) e AR(1) mal specificati. Per modello correttamente specificato si intende che la serie storica è stata generata da quel processo. Questi tre esperimenti Monte Carlo saranno spiegati in modo più dettagliato nelle relative sezioni.

Ogni simulazione è svolta generando serie storiche indipendenti con una certa numerosità campionaria n e calcolando le previsioni per $h = 1$ e 2 passi in avanti. Nei casi in cui le previsioni ottenute non siano discrete, ciò può accadere ad esempio quando si prevede col modello AR(1), queste sono state

arrotondate all'intero più vicino, in modo tale da poter confrontare previsioni coerenti con la caratteristica di conteggio del processo. La qualità delle previsioni è valutata attraverso diversi criteri; in particolare, sono calcolati tre indicatori per le previsioni puntuali e due indicatori per le previsioni dell'intera distribuzione di probabilità. I criteri sono i seguenti:

1. Errore assoluto medio di previsione (MAE), per previsioni ad h passi in avanti, calcolato con la formula: $\frac{1}{N} \sum_{i=1}^N |\hat{x}_{t+h}^i - x_{t+h}^i|$ dove, \hat{x}_{t+h}^i è la previsione ad h passi dell' i -esima serie simulata e N è il numero di simulazioni effettuate.
2. Errore quadratico medio di previsione (MSE), per previsioni ad h passi in avanti, calcolato con la formula: $\frac{1}{N} \sum_{i=1}^N (\hat{x}_{t+h}^i - x_{t+h}^i)^2$.
3. Frequenza di corretta previsione, per previsioni ad h passi in avanti, calcolata con la formula: $\frac{1}{N} \sum_{i=1}^N I(\hat{x}_{t+h}^i = x_{t+h}^i)$ dove, $I(\hat{x}_{t+h}^i = x_{t+h}^i)$ è una funzione indicatrice che è pari a 1 se $\hat{x}_{t+h}^i = x_{t+h}^i$ e zero altrimenti. Maggiore è il valore assunto da questo indicatore e migliore è la qualità delle previsioni dato che rappresenta la frequenza con cui la previsione è esatta.
4. La "*total variation distance*" tra la vera distribuzione condizionale ad h passi e la distribuzione prevista, calcolata con la formula:

$$\frac{1}{N} \sum_{i=1}^N \sum_{x=0}^{\infty} \frac{|\hat{p}(x_{t+h}^i = x) - p(x_{t+h}^i = x)|}{2}$$
dove, $\hat{p}(x_{t+h}^i = x)$ è la previsione della funzione di probabilità ad h passi per l' i -esima serie simulata. Questo indicatore è compreso tra zero e uno, e valori elevati indicano che la previsione non è buona dato che rappresenta una distanza tra le due distribuzioni.
5. La distanza di Bhattacharyya, Bhattacharyya (1986), tra la vera distribuzione condizionale ad h passi in avanti e la distribuzione prevista, calcolata con la formula: $\frac{1}{N} \sum_{i=1}^N -\log \left(\sum_{x=0}^{\infty} \hat{p}(x_{t+h}^i = x) p(x_{t+h}^i = x) \right)$.

Questo indicatore è sempre maggiore uguale di zero, e valori elevati indicano che la previsione non è buona dato che rappresenta una distanza tra le due distribuzioni.

3.3.2 Innovazione Poisson

Questo primo esperimento è un confronto, secondo i criteri esposti in precedenza, tra le previsioni puntuali ottenute con il modello PoINAR(1) correttamente specificato e quelle ottenute da un classico AR(1) mal specificato. Le osservazioni sono generate dal processo PoINAR(1) e sono considerati diversi valori dei parametri. In particolare, il fine di questo esperimento è verificare se l'utilizzo del classico modello autoregressivo su dati di conteggio possa essere un valido metodo per ottenere previsioni puntuali e rendere così poco utile l'utilizzo di un modello correttamente specificato per dati di conteggio.

I risultati delle simulazioni sono presentati nelle tabelle 3.1, 3.2 e 3.3. Sulla base di questi risultati si può affermare che l'utilizzo del modello PoINAR(1) non porti a sensibili benefici rispetto all'AR(1). Le differenze in termini di *performance* sono minime e non sembrano esserci particolari effetti degni di nota né al variare della numerosità campionaria né al variare dei valori assunti dai parametri. Si può quindi concludere che, il modello AR(1) può essere tranquillamente utilizzato come alternativa al PoINAR(1) per ottenere previsioni puntuali da serie storiche di conteggio. Ovviamente, se l'obiettivo è quello di ricavare previsioni dell'intera distribuzione di probabilità, l'AR(1) non è utilizzabile ed è necessario adottare un modello per dati di conteggio come il PoINAR(1).

Tabella 3.1: Errore assoluto medio di previsione (FMAE), per previsioni h passi in avanti, calcolato su $N=1000$ simulazioni per dati generati dal processo INAR(1) con innovazione Poisson. I risultati migliori sono evidenziati in grassetto.

Parametri	Modello	Previsione	$n=50$		$n=100$		$n=300$	
			$h = 1$	$h = 2$	$h = 1$	$h = 2$	$h = 1$	$h = 2$
$\alpha = 0.2$ $\lambda = 1$	Poisson	Mediana	0.855	0.820	0.782	0.799	0.835	0.828
		Moda	0.905	0.856	0.820	0.816	0.872	0.828
		Media	0.853	0.827	0.774	0.806	0.855	0.832
	AR(1)		0.855	0.830	0.778	0.809	0.850	0.832
$\alpha = 0.5$ $\lambda = 1$	Poisson	Mediana	0.954	1.033	0.929	1.051	0.991	1.073
		Moda	0.984	1.081	0.952	1.072	1.001	1.119
		Media	0.947	1.021	0.935	1.056	1.002	1.083
	AR(1)		0.942	1.029	0.939	1.047	1.008	1.085
$\alpha = 0.8$ $\lambda = 1$	Poisson	Mediana	1.076	1.399	1.017	1.320	0.981	1.345
		Moda	1.063	1.410	1.006	1.327	0.995	1.360
		Media	1.073	1.409	1.011	1.335	0.981	1.345
	AR(1)		1.073	1.416	1.024	1.344	0.986	1.352
$\alpha = 0.2$ $\lambda = 3$	Poisson	Mediana	1.541	1.529	1.527	1.547	1.459	1.524
		Moda	1.566	1.552	1.580	1.540	1.483	1.556
		Media	1.555	1.528	1.528	1.573	1.495	1.521
	AR(1)		1.548	1.537	1.531	1.572	1.495	1.522
$\alpha = 0.5$ $\lambda = 3$	Poisson	Mediana	1.696	1.963	1.744	1.857	1.681	1.847
		Moda	1.721	1.971	1.737	1.852	1.673	1.876
		Media	1.716	1.955	1.749	1.881	1.672	1.861
	AR(1)		1.706	1.964	1.743	1.884	1.669	1.856
$\alpha = 0.8$ $\lambda = 3$	Poisson	Mediana	1.825	2.335	1.934	2.485	1.822	2.354
		Moda	1.834	2.362	1.939	2.494	1.813	2.362
		Media	1.822	2.335	1.926	2.492	1.819	2.370
	AR(1)		1.840	2.372	1.933	2.491	1.833	2.386

Tabella 3.2: Errore quadratico medio di previsione (FMSE), per previsioni h passi in avanti, calcolato su $N=1000$ simulazioni per dati generati dal processo INAR(1) con innovazione Poisson. I risultati migliori sono evidenziati in grassetto.

Parametri	Modello	Previsione	$n=50$		$n=100$		$n=300$	
			$h = 1$	$h = 2$	$h = 1$	$h = 2$	$h = 1$	$h = 2$
$\alpha = 0.2$ $\lambda = 1$	Poisson	Mediana	1.375	1.232	1.264	1.239	1.289	1.326
		Moda	1.603	1.374	1.440	1.300	1.442	1.326
		Media	1.371	1.259	1.206	1.252	1.317	1.338
	AR(1)		1.375	1.264	1.220	1.259	1.300	1.334
$\alpha = 0.5$ $\lambda = 1$	Poisson	Mediana	1.590	1.995	1.587	2.035	1.789	2.135
		Moda	1.684	2.185	1.648	2.154	1.863	2.377
		Media	1.559	1.897	1.587	2.024	1.789	2.117
	AR(1)		1.584	1.933	1.609	1.987	1.816	2.113
$\alpha = 0.8$ $\lambda = 1$	Poisson	Mediana	2.008	3.267	1.901	2.988	1.783	3.017
		Moda	2.003	3.400	1.910	3.085	1.813	3.050
		Media	2.005	3.269	1.891	3.003	1.771	3.035
	AR(1)		2.039	3.322	1.896	3.040	1.782	3.052
$\alpha = 0.2$ $\lambda = 3$	Poisson	Mediana	3.949	3.879	3.817	4.015	3.601	3.808
		Moda	4.174	4.044	4.112	4.238	3.801	4.144
		Media	3.975	3.804	3.782	3.993	3.689	3.709
	AR(1)		3.970	3.823	3.761	4.000	3.663	3.718
$\alpha = 0.5$ $\lambda = 3$	Poisson	Mediana	4.944	6.483	4.862	5.835	4.537	5.613
		Moda	5.101	6.655	4.973	5.960	4.575	5.860
		Media	4.988	6.409	4.897	5.865	4.492	5.613
	AR(1)		4.948	6.466	4.903	5.866	4.485	5.612
$\alpha = 0.8$ $\lambda = 3$	Poisson	Mediana	5.575	9.373	6.128	9.973	5.534	8.858
		Moda	5.616	9.522	6.175	10.004	5.529	8.930
		Media	5.578	9.343	6.100	9.994	5.527	8.988
	AR(1)		5.646	9.552	6.121	10.053	5.601	9.078

Tabella 3.3: Frequenza di corretta previsione, per previsioni h passi in avanti, calcolato su $N=1000$ simulazioni per dati generati dal processo INAR(1) con innovazione Poisson. I risultati migliori sono evidenziati in grassetto.

Parametri	Modello	Previsione	$n=50$		$n=100$		$n=300$	
			$h = 1$	$h = 2$	$h = 1$	$h = 2$	$h = 1$	$h = 2$
$\alpha = 0.2$ $\lambda = 1$	Poisson	Mediana	0.344	0.347	0.393	0.365	0.346	0.360
		Moda	0.347	0.346	0.406	0.365	0.343	0.360
		Media	0.348	0.351	0.388	0.364	0.335	0.360
	AR(1)		0.347	0.349	0.386	0.363	0.336	0.359
$\alpha = 0.5$ $\lambda = 1$	Poisson	Mediana	0.310	0.316	0.321	0.299	0.307	0.309
		Moda	0.296	0.297	0.313	0.285	0.314	0.301
		Media	0.315	0.311	0.321	0.303	0.300	0.297
	AR(1)		0.327	0.313	0.324	0.304	0.298	0.295
$\alpha = 0.8$ $\lambda = 1$	Poisson	Mediana	0.284	0.221	0.315	0.233	0.336	0.230
		Moda	0.300	0.234	0.329	0.247	0.328	0.218
		Media	0.288	0.212	0.321	0.221	0.333	0.234
	AR(1)		0.298	0.212	0.310	0.216	0.332	0.228
$\alpha = 0.2$ $\lambda = 3$	Poisson	Mediana	0.192	0.208	0.200	0.194	0.218	0.204
		Moda	0.200	0.210	0.191	0.217	0.220	0.200
		Media	0.192	0.198	0.197	0.178	0.204	0.198
	AR(1)		0.194	0.193	0.194	0.179	0.197	0.197
$\alpha = 0.5$ $\lambda = 3$	Poisson	Mediana	0.195	0.169	0.169	0.177	0.175	0.172
		Moda	0.196	0.166	0.176	0.191	0.186	0.177
		Media	0.189	0.156	0.169	0.171	0.180	0.162
	AR(1)		0.191	0.166	0.172	0.172	0.182	0.165
$\alpha = 0.8$ $\lambda = 3$	Poisson	Mediana	0.187	0.144	0.168	0.136	0.177	0.131
		Moda	0.180	0.131	0.165	0.125	0.182	0.131
		Media	0.191	0.145	0.170	0.135	0.179	0.135
	AR(1)		0.181	0.143	0.173	0.131	0.177	0.130

3.3.3 Innovazione Binomiale

In questa sezione è svolto un confronto tra le previsioni ottenute con il modello INAR(1) con errore Binomiale correttamente specificato e quelle ottenute dai modelli PoINAR(1) e AR(1) mal specificati. Le osservazioni sono generate dal processo Binomiale e i valori dei parametri sono stati scelti in modo tale da distinguere due casi. Il primo caso ($m=4$ e $p=0.8$) impone che la sottodispersione del processo sia molto marcata ($\mu_\epsilon = 3.2$ e $\sigma_\epsilon^2 = 0.64$), creando quindi una situazione poco adatta al modello Poisson. Il secondo caso ($m=16$ e $p=0.2$) impone un minor grado di sottodispersione ($\mu_\epsilon = 3.2$ e $\sigma_\epsilon^2 = 2.56$), creando così una situazione più congeniale al modello Poisson.

Lo scopo di questo esperimento è verificare se l'utilizzo del modello Binomiale possa portare dei vantaggi in termini di qualità delle previsioni.

I risultati delle simulazioni per le previsioni puntuali sono presentati nelle tabelle 3.4, 3.5 e 3.6, mentre quelli per le previsioni della distribuzione di probabilità sono esposti in tabella 3.7.

Per quanto riguarda le previsioni puntuali, le simulazioni mostrano che, le prestazioni dei tre modelli sono sostanzialmente equivalenti quando la sottodispersione è meno evidente; mentre, quando la sottodispersione è più marcata, il modello Poisson ha *performance* inferiori rispetto al Binomiale e all'AR(1). Queste differenze sembrano comunque attenuarsi per valori elevati di α . Relativamente alle previsioni della distribuzione di probabilità i risultati mostrano che, il modello Binomiale porta a risultati decisamente migliori nel caso di elevata sottodispersione; mentre, quando la sottodispersione è meno accentuata, le *performance* dei due modelli sono molto simili soprattutto a basse numerosità campionarie.

Si può quindi concludere che, il modello Binomiale può essere considerato utile limitatamente al caso in cui lo scopo della modellazione sia quello di prevedere delle probabilità condizionali e sia inoltre presente un evidente sottodispersione nelle osservazioni.

Tabella 3.4: Errore assoluto medio di previsione (FMAE), per previsioni h passi in avanti, calcolato su $N=1000$ simulazioni per dati generati dal processo INAR(1) con innovazione Binomiale. I risultati migliori sono evidenziati in grassetto.

Parametri	Modello	Previsione	$n=50$		$n=100$		$n=300$	
			$h = 1$	$h = 2$	$h = 1$	$h = 2$	$h = 1$	$h = 2$
$\alpha = 0.2$ $m = 4$ $p = 0.8$	Binomiale	Mediana	0.849	0.856	0.840	0.854	0.813	0.835
		Media	0.852	0.859	0.835	0.856	0.815	0.835
	Poisson	Mediana	1.041	1.004	1.030	0.973	0.978	1.017
		Media	1.038	1.016	1.011	0.973	0.951	0.986
	AR		0.855	0.858	0.845	0.853	0.811	0.835
$\alpha = 0.5$ $m = 4$ $p = 0.8$	Binomiale	Mediana	1.232	1.390	1.153	1.356	1.190	1.308
		Media	1.234	1.390	1.157	1.347	1.191	1.301
	Poisson	Mediana	1.289	1.443	1.191	1.401	1.268	1.389
		Media	1.284	1.442	1.190	1.393	1.252	1.352
	AR		1.224	1.389	1.152	1.351	1.187	1.303
$\alpha = 0.8$ $m = 4$ $p = 0.8$	Binomiale	Mediana	1.444	1.841	1.432	1.826	1.441	1.871
		Media	1.462	1.831	1.419	1.818	1.434	1.865
	Poisson	Mediana	1.463	1.842	1.453	1.824	1.445	1.894
		Media	1.455	1.840	1.437	1.811	1.446	1.902
	AR		1.471	1.854	1.413	1.816	1.422	1.876
$\alpha = 0.2$ $m = 16$ $p = 0.2$	Binomiale	Mediana	1.365	1.485	1.360	1.452	1.442	1.382
		Media	1.366	1.475	1.356	1.450	1.444	1.378
	Poisson	Mediana	1.365	1.497	1.372	1.475	1.449	1.384
		Media	1.365	1.479	1.368	1.461	1.452	1.385
	AR		1.369	1.475	1.352	1.448	1.448	1.379
$\alpha = 0.5$ $m = 16$ $p = 0.2$	Binomiale	Mediana	1.573	1.876	1.617	1.831	1.648	1.793
		Media	1.579	1.864	1.626	1.836	1.646	1.804
	Poisson	Mediana	1.575	1.863	1.621	1.846	1.644	1.795
		Media	1.569	1.846	1.635	1.846	1.652	1.797
	AR		1.572	1.871	1.634	1.839	1.652	1.800
$\alpha = 0.8$ $m = 16$ $p = 0.2$	Binomiale	Mediana	1.869	2.396	1.826	2.370	1.749	2.262
		Media	1.875	2.389	1.823	2.374	1.749	2.266
	Poisson	Mediana	1.843	2.345	1.807	2.365	1.747	2.266
		Media	1.846	2.357	1.813	2.354	1.740	2.269
	AR		1.875	2.386	1.839	2.380	1.752	2.257

Tabella 3.5: Errore quadratico medio di previsione (FMSE), per previsioni h passi in avanti, calcolato su $N=1000$ simulazioni per dati generati dal processo INAR(1) con innovazione Binomiale. I risultati migliori sono evidenziati in grassetto.

Parametri	Modello	Previsione	$n=50$		$n=100$		$n=300$	
			$h = 1$	$h = 2$	$h = 1$	$h = 2$	$h = 1$	$h = 2$
$\alpha = 0.2$ $m = 4$ $p = 0.8$	Binomiale	Mediana	1.343	1.344	1.316	1.346	1.253	1.317
		Media	1.346	1.355	1.308	1.350	1.261	1.317
	Poisson	Mediana	1.841	1.824	1.752	1.711	1.640	1.859
		Media	1.834	1.844	1.713	1.707	1.571	1.752
	AR		1.357	1.344	1.329	1.347	1.233	1.317
$\alpha = 0.5$ $m = 4$ $p = 0.8$	Binomiale	Mediana	2.518	3.062	2.299	2.966	2.402	2.852
		Media	2.528	3.066	2.311	2.949	2.401	2.835
	Poisson	Mediana	2.753	3.329	2.445	3.235	2.586	3.205
		Media	2.760	3.316	2.432	3.207	2.526	3.076
	AR		2.524	3.047	2.292	2.951	2.387	2.835
$\alpha = 0.8$ $m = 4$ $p = 0.8$	Binomiale	Mediana	3.408	5.535	3.374	5.394	3.397	5.693
		Media	3.406	5.513	3.321	5.378	3.324	5.665
	Poisson	Mediana	3.415	5.574	3.463	5.406	3.409	5.818
		Media	3.401	5.578	3.419	5.363	3.418	5.862
	AR		3.455	5.602	3.345	5.424	3.302	5.682
$\alpha = 0.2$ $m = 16$ $p = 0.2$	Binomiale	Mediana	3.163	3.725	3.136	3.404	3.458	3.134
		Media	3.148	3.605	3.110	3.362	3.422	3.128
	Poisson	Mediana	3.143	3.757	3.198	3.501	3.475	3.150
		Media	3.133	3.645	3.172	3.411	3.434	3.155
	AR		3.139	3.613	3.112	3.360	3.426	3.129
$\alpha = 0.5$ $m = 16$ $p = 0.2$	Binomiale	Mediana	4.143	5.638	4.183	5.451	4.370	5.163
		Media	4.169	5.536	4.196	5.454	4.350	5.226
	Poisson	Mediana	4.191	5.579	4.181	5.520	4.384	5.221
		Media	4.165	5.476	4.209	5.514	4.374	5.193
	AR		4.178	5.625	4.220	5.469	4.370	5.210
$\alpha = 0.8$ $m = 16$ $p = 0.2$	Binomiale	Mediana	5.541	9.140	5.466	9.032	4.907	8.366
		Media	5.557	9.107	5.403	9.052	4.909	8.370
	Poisson	Mediana	5.449	8.817	5.423	9.002	4.925	8.388
		Media	5.470	8.855	5.431	8.926	4.880	8.393
	AR		5.603	9.074	5.513	9.122	4.946	8.337

Tabella 3.6: Frequenza di corretta previsione, per previsioni h passi in avanti, calcolata su $N=1000$ simulazioni per dati generati dal processo INAR(1) con innovazione Binomiale. I risultati migliori sono evidenziati in grassetto.

Parametri	Modello	Previsione	$n=50$		$n=100$		$n=300$	
			$h = 1$	$h = 2$	$h = 1$	$h = 2$	$h = 1$	$h = 2$
$\alpha = 0.2$ $m = 4$ $p = 0.8$	Binomiale	Mediana	0.366	0.356	0.362	0.364	0.373	0.365
		Media	0.365	0.357	0.361	0.363	0.373	0.365
	Poisson	Mediana	0.279	0.325	0.275	0.328	0.297	0.307
		Media	0.276	0.314	0.283	0.327	0.307	0.308
	AR		0.363	0.354	0.361	0.366	0.366	0.365
$\alpha = 0.5$ $m = 4$ $p = 0.8$	Binomiale	Mediana	0.252	0.207	0.271	0.219	0.261	0.233
		Media	0.253	0.207	0.269	0.224	0.259	0.235
	Poisson	Mediana	0.239	0.207	0.259	0.221	0.223	0.222
		Media	0.245	0.203	0.258	0.223	0.226	0.237
	AR		0.262	0.207	0.270	0.219	0.262	0.234
$\alpha = 0.8$ $m = 4$ $p = 0.8$	Binomiale	Mediana	0.209	0.168	0.226	0.171	0.222	0.163
		Media	0.196	0.171	0.232	0.179	0.216	0.163
	Poisson	Mediana	0.191	0.168	0.219	0.175	0.222	0.170
		Media	0.196	0.172	0.227	0.178	0.222	0.165
	AR		0.191	0.174	0.238	0.175	0.220	0.157
$\alpha = 0.2$ $m = 16$ $p = 0.2$	Binomiale	Mediana	0.243	0.216	0.224	0.201	0.221	0.212
		Media	0.236	0.205	0.229	0.194	0.212	0.214
	Poisson	Mediana	0.239	0.209	0.224	0.194	0.209	0.210
		Media	0.233	0.208	0.223	0.193	0.205	0.213
	AR		0.232	0.207	0.233	0.196	0.207	0.213
$\alpha = 0.5$ $m = 16$ $p = 0.2$	Binomiale	Mediana	0.202	0.162	0.184	0.179	0.189	0.159
		Media	0.203	0.165	0.181	0.178	0.190	0.156
	Poisson	Mediana	0.200	0.170	0.178	0.171	0.191	0.167
		Media	0.201	0.167	0.173	0.168	0.188	0.165
	AR		0.212	0.169	0.180	0.178	0.187	0.160
$\alpha = 0.8$ $m = 16$ $p = 0.2$	Binomiale	Mediana	0.159	0.118	0.184	0.138	0.174	0.152
		Media	0.160	0.119	0.178	0.138	0.173	0.152
	Poisson	Mediana	0.164	0.126	0.189	0.132	0.180	0.143
		Media	0.167	0.117	0.183	0.130	0.178	0.143
	AR		0.165	0.120	0.180	0.136	0.176	0.155

Tabella 3.7: *Total variation distance* e *Bhattacharyya distance* tra la vera distribuzione h passi in avanti e quella prevista, calcolate su $N=1000$ simulazioni per dati generati dal processo INAR(1) con innovazione Binomiale. I risultati migliori sono evidenziati in grassetto.

<i>Total variation distance</i>							
Parametri	Modello	$n=50$		$n=100$		$n=300$	
		$h = 1$	$h = 2$	$h = 1$	$h = 2$	$h = 1$	$h = 2$
$\alpha = 0.2, m = 4$ $p = 0.8$	Binomiale	0.090	0.089	0.061	0.060	0.033	0.033
	Poisson	0.206	0.244	0.194	0.238	0.202	0.242
$\alpha = 0.5, m = 4$ $p = 0.8$	Binomiale	0.082	0.092	0.054	0.062	0.030	0.036
	Poisson	0.129	0.159	0.122	0.154	0.118	0.151
$\alpha = 0.8, m = 4$ $p = 0.8$	Binomiale	0.078	0.093	0.050	0.061	0.029	0.038
	Poisson	0.103	0.111	0.090	0.096	0.079	0.083
$\alpha = 0.2, m = 16$ $p = 0.2$	Binomiale	0.078	0.074	0.053	0.050	0.031	0.029
	Poisson	0.080	0.074	0.065	0.061	0.051	0.051
$\alpha = 0.5, m = 16$ $p = 0.2$	Binomiale	0.083	0.091	0.056	0.062	0.031	0.035
	Poisson	0.072	0.078	0.056	0.061	0.040	0.045
$\alpha = 0.8, m = 16$ $p = 0.2$	Binomiale	0.084	0.103	0.057	0.069	0.030	0.035
	Poisson	0.069	0.081	0.051	0.059	0.033	0.036

<i>Bhattacharyya distance</i>							
Parametri	Modello	$n=50$		$n=100$		$n=300$	
		$h = 1$	$h = 2$	$h = 1$	$h = 2$	$h = 1$	$h = 2$
$\alpha = 0.2, m = 4$ $p = 0.8$	Binomiale	0.0105	0.0090	0.0048	0.0041	0.0015	0.0012
	Poisson	0.0470	0.0604	0.0422	0.0579	0.0407	0.0574
$\alpha = 0.5, m = 4$ $p = 0.8$	Binomiale	0.0117	0.0120	0.0053	0.0053	0.0013	0.0016
	Poisson	0.0237	0.0300	0.0205	0.0267	0.0185	0.0247
$\alpha = 0.8, m = 4$ $p = 0.8$	Binomiale	0.0094	0.0106	0.0045	0.0052	0.0013	0.0019
	Poisson	0.0151	0.0147	0.0129	0.0116	0.0111	0.0091
$\alpha = 0.2, m = 16$ $p = 0.2$	Binomiale	0.0074	0.0065	0.0037	0.0031	0.0011	0.0010
	Poisson	0.0070	0.0059	0.0047	0.0040	0.0029	0.0028
$\alpha = 0.5, m = 16$ $p = 0.2$	Binomiale	0.0084	0.0097	0.0041	0.0047	0.0012	0.0014
	Poisson	0.0060	0.0068	0.0034	0.0039	0.0017	0.0021
$\alpha = 0.8, m = 16$ $p = 0.2$	Binomiale	0.0095	0.0137	0.0041	0.0058	0.0012	0.0016
	Poisson	0.0058	0.0079	0.0030	0.0040	0.0014	0.0017

3.3.4 Innovazione Binomiale negativa

In questo terzo ed ultimo esperimento, è svolto un confronto tra le previsioni ottenute con il modello INAR(1) con errore Binomiale Negativa correttamente specificato e quelle ottenute dai modelli PoINAR(1) a AR(1) mal specificati. Le osservazioni sono generate dal processo INAR(1) con innovazione Binomiale Negativa e i valori dei parametri sono stati scelti in modo tale da distinguere due casi. Il primo caso ($r = 1.3$ e $p = 0.3$) impone un'elevata sovradisersione ($\mu_\epsilon \approx 3$ e $\sigma_\epsilon^2 \approx 10$), creando quindi una situazione poco adatta al modello Poisson. Il secondo caso ($r = 12$ e $p = 0.8$) impone una sottodispersione poco marcata ($\mu_\epsilon = 3$ e $\sigma_\epsilon^2 = 3.75$), creando così una situazione più congeniale al modello Poisson.

Lo scopo di questo esperimento è verificare se l'utilizzo del modello Binomiale Negativa possa portare dei vantaggi in termini di qualità delle previsioni.

I risultati delle simulazioni per le previsioni puntuali sono presentati nelle tabelle 3.8, 3.9 e 3.10, mentre quelli per le previsioni della distribuzione di probabilità sono esposti in tabella 3.11.

Per quanto riguarda le previsioni puntuali le simulazioni mostrano che, nel caso di elevata sovradisersione, le prestazioni del modello correttamente specificato sembrano leggermente migliori ma le differenze sono comunque minime e da considerarsi sostanzialmente trascurabili; mentre, nel caso in cui la sovradisersione è meno accentuata, le prestazioni dei tre modelli risultano equivalenti. Relativamente alle previsioni dell'intera distribuzione di probabilità si nota che, il modello con la Binomiale Negativa porta a risultati decisamente migliori nel caso di elevata sovradisersione; invece, quando la sovradisersione è meno accentuata, le *performance* dei due modelli sono molto simili soprattutto a basse numerosità campionarie. Si può quindi concludere che, il modello Binomiale Negativa può essere considerato utile limitatamente al caso in cui lo scopo della modellazione sia quello di prevedere delle probabilità condizionali e sia, inoltre, presente un'evidente sovradisersione nelle osservazioni.

Tabella 3.8: Errore assoluto medio di previsione (FMAE), per previsioni h passi in avanti, calcolato su $N=1000$ simulazioni per dati generati dal processo INAR(1) con innovazione Binomiale Negativa. I risultati migliori sono evidenziati in grassetto.

Parametri	Modello	Previsione	$n=50$		$n=100$		$n=300$	
			$h = 1$	$h = 2$	$h = 1$	$h = 2$	$h = 1$	$h = 2$
$\alpha = 0.2$ $r = 1.3$ $p = 0.3$	B. Negativa	Mediana	2.303	2.452	2.393	2.328	2.441	2.246
		Media	2.451	2.570	2.423	2.457	2.519	2.382
	Poisson	Mediana	2.448	2.534	2.418	2.408	2.504	2.346
		Media	2.490	2.568	2.445	2.452	2.527	2.389
	AR		2.485	2.578	2.434	2.450	2.507	2.376
$\alpha = 0.5$ $r = 1.3$ $p = 0.3$	B. Negativa	Mediana	2.475	2.821	2.532	2.854	2.370	2.873
		Media	2.538	2.855	2.587	2.912	2.468	2.967
	Poisson	Mediana	2.571	2.837	2.608	2.927	2.483	2.926
		Media	2.598	2.880	2.630	2.959	2.509	2.970
	AR		2.553	2.869	2.591	2.924	2.478	2.967
$\alpha = 0.8$ $r = 1.3$ $p = 0.3$	B. Negativa	Mediana	2.674	3.580	2.698	3.565	2.612	3.461
		Media	2.747	3.644	2.764	3.638	2.662	3.474
	Poisson	Mediana	2.819	3.820	2.865	3.778	2.735	3.649
		Media	2.820	3.813	2.882	3.793	2.753	3.649
	AR		2.753	3.689	2.786	3.645	2.637	3.474
$\alpha = 0.2$ $r = 12$ $p = 0.8$	B. Negativa	Mediana	1.656	1.649	1.676	1.691	1.655	1.729
		Media	1.663	1.690	1.724	1.708	1.664	1.709
	Poisson	Mediana	1.646	1.656	1.691	1.690	1.647	1.719
		Media	1.651	1.682	1.724	1.711	1.660	1.705
	AR		1.654	1.680	1.727	1.708	1.664	1.711
$\alpha = 0.5$ $r = 12$ $p = 0.8$	B. Negativa	Mediana	1.768	2.047	1.802	2.131	1.831	2.036
		Media	1.775	2.042	1.799	2.151	1.842	2.057
	Poisson	Mediana	1.785	2.045	1.799	2.146	1.830	2.030
		Media	1.782	2.048	1.812	2.155	1.840	2.049
	AR		1.788	2.034	1.806	2.158	1.841	2.055
$\alpha = 0.8$ $r = 12$ $p = 0.8$	B. Negativa	Mediana	1.896	2.590	1.936	2.548	1.948	2.484
		Media	1.898	2.587	1.923	2.564	1.954	2.494
	Poisson	Mediana	1.891	2.582	1.932	2.550	1.956	2.496
		Media	1.895	2.575	1.924	2.563	1.955	2.494
	AR		1.926	2.617	1.927	2.581	1.963	2.491

Tabella 3.9: Errore quadratico medio di previsione (FMSE), per previsioni h passi in avanti, calcolato su $N=1000$ simulazioni per dati generati dal processo INAR(1) con innovazione Binomiale Negativa. I risultati migliori sono evidenziati in grassetto.

Parametri	Modello	Previsione	$n=50$		$n=100$		$n=300$	
			$h = 1$	$h = 2$	$h = 1$	$h = 2$	$h = 1$	$h = 2$
$\alpha = 0.2$ $r = 1.3$ $p = 0.3$	B. Negativa	Mediana	10.86	11.67	11.59	11.21	12.88	10.37
		Media	10.49	11.37	10.58	10.74	12.14	9.99
	Poisson	Mediana	10.67	11.38	10.80	10.64	12.31	10.00
		Media	10.77	11.40	10.74	10.68	12.25	9.97
	AR		10.80	11.50	10.71	10.72	12.07	9.95
$\alpha = 0.5$ $r = 1.3$ $p = 0.3$	B. Negativa	Mediana	11.99	14.66	12.53	14.88	11.71	15.41
		Media	11.48	14.18	12.00	14.31	11.27	14.90
	Poisson	Mediana	11.91	14.28	12.53	14.88	11.44	14.94
		Media	11.87	14.34	12.26	14.44	11.41	15.04
	AR		11.57	14.30	12.02	14.43	11.31	14.89
$\alpha = 0.8$ $r = 1.3$ $p = 0.3$	B. Negativa	Mediana	13.26	22.54	14.24	22.02	13.30	20.45
		Media	13.19	22.37	13.84	22.11	12.55	19.94
	Poisson	Mediana	13.88	24.41	14.62	23.70	13.39	21.58
		Media	13.91	24.33	14.62	23.78	13.37	21.56
	AR		13.34	22.99	14.00	22.26	12.62	19.95
$\alpha = 0.2$ $r = 12$ $p = 0.8$	B. Negativa	Mediana	4.61	4.54	4.68	4.88	4.57	5.12
		Media	4.52	4.58	4.72	4.73	4.48	4.75
	Poisson	Mediana	4.54	4.52	4.68	4.79	4.50	4.96
		Media	4.49	4.56	4.73	4.73	4.45	4.73
	AR		4.52	4.57	4.74	4.73	4.46	4.75
$\alpha = 0.5$ $r = 12$ $p = 0.8$	B. Negativa	Mediana	5.15	6.77	5.37	7.53	5.35	6.90
		Media	5.12	6.72	5.36	7.47	5.35	6.93
	Poisson	Mediana	5.18	6.71	5.38	7.54	5.37	6.82
		Media	5.14	6.73	5.38	7.51	5.36	6.90
	AR		5.16	6.64	5.39	7.54	5.34	6.91
$\alpha = 0.8$ $r = 12$ $p = 0.8$	B. Negativa	Mediana	5.89	10.93	6.15	10.79	6.27	10.12
		Media	5.86	10.86	6.09	10.75	6.27	10.17
	Poisson	Mediana	5.84	10.78	6.11	10.77	6.30	10.13
		Media	5.83	10.71	6.08	10.78	6.28	10.12
	AR		6.02	11.11	6.09	10.86	6.39	10.15

Tabella 3.10: Frequenza di corretta previsione, per previsioni h passi in avanti, calcolata su $N=1000$ simulazioni per dati generati dal processo INAR(1) con innovazione Binomiale Negativa. I risultati migliori sono evidenziati in grassetto.

Parametri	Modello	Previsione	$n=50$		$n=100$		$n=300$	
			$h = 1$	$h = 2$	$h = 1$	$h = 2$	$h = 1$	$h = 2$
$\alpha = 0.2$ $r = 1.3$ $p = 0.3$	B. Negativa	Mediana	0.161	0.129	0.145	0.153	0.130	0.158
		Media	0.114	0.116	0.127	0.115	0.128	0.119
	Poisson	Mediana	0.115	0.119	0.128	0.122	0.125	0.124
		Media	0.105	0.119	0.119	0.113	0.131	0.120
	AR		0.108	0.121	0.123	0.118	0.130	0.121
$\alpha = 0.5$ $r = 1.3$ $p = 0.3$	B. Negativa	Mediana	0.131	0.117	0.155	0.132	0.176	0.122
		Media	0.124	0.117	0.122	0.103	0.124	0.092
	Poisson	Mediana	0.114	0.119	0.122	0.110	0.134	0.110
		Media	0.104	0.116	0.116	0.096	0.123	0.096
	AR		0.120	0.118	0.116	0.105	0.119	0.090
$\alpha = 0.8$ $r = 1.3$ $p = 0.3$	B. Negativa	Mediana	0.117	0.105	0.144	0.090	0.138	0.082
		Media	0.103	0.092	0.120	0.085	0.131	0.091
	Poisson	Mediana	0.107	0.085	0.105	0.087	0.112	0.071
		Media	0.112	0.088	0.099	0.091	0.115	0.082
	AR		0.106	0.093	0.111	0.090	0.123	0.091
$\alpha = 0.2$ $r = 12$ $p = 0.8$	B. Negativa	Mediana	0.202	0.185	0.181	0.183	0.185	0.182
		Media	0.186	0.169	0.159	0.165	0.170	0.181
	Poisson	Mediana	0.201	0.178	0.171	0.177	0.183	0.175
		Media	0.194	0.172	0.158	0.167	0.168	0.184
	AR		0.192	0.176	0.158	0.164	0.168	0.178
$\alpha = 0.5$ $r = 12$ $p = 0.8$	B. Negativa	Mediana	0.178	0.151	0.170	0.152	0.160	0.157
		Media	0.166	0.152	0.178	0.137	0.162	0.141
	Poisson	Mediana	0.167	0.151	0.170	0.146	0.165	0.147
		Media	0.160	0.151	0.166	0.141	0.163	0.148
	AR		0.160	0.148	0.173	0.140	0.160	0.145
$\alpha = 0.8$ $r = 12$ $p = 0.8$	B. Negativa	Mediana	0.170	0.124	0.161	0.150	0.149	0.133
		Media	0.162	0.128	0.166	0.134	0.147	0.134
	Poisson	Mediana	0.162	0.128	0.164	0.149	0.147	0.123
		Media	0.163	0.131	0.171	0.142	0.149	0.127
	AR		0.157	0.128	0.165	0.137	0.154	0.133

Tabella 3.11: *Total variation distance* e *Bhattacharyya distance* tra la vera distribuzione ad h passi in avanti e quella prevista, calcolate su $N=1000$ simulazioni per dati generati dal processo INAR(1) con innovazione Binomiale Negativa. I risultati migliori sono evidenziati in grassetto.

<i>Total variation distance</i>							
Parametri	Modello	$n=50$		$n=100$		$n=300$	
		$h = 1$	$h = 2$	$h = 1$	$h = 2$	$h = 1$	$h = 2$
$\alpha = 0.2, r = 1.3$ $p = 0.3$	B. Negativa	0.085	0.080	0.060	0.056	0.034	0.031
	Poisson	0.262	0.261	0.261	0.259	0.257	0.254
$\alpha = 0.5, r = 1.3$ $p = 0.3$	B. Negativa	0.087	0.091	0.060	0.064	0.033	0.036
	Poisson	0.217	0.231	0.209	0.223	0.203	0.218
$\alpha = 0.8, r = 1.3$ $p = 0.3$	B. Negativa	0.086	0.096	0.061	0.067	0.034	0.037
	Poisson	0.171	0.185	0.165	0.177	0.161	0.170
$\alpha = 0.2, r = 12$ $p = 0.8$	B. Negativa	0.076	0.071	0.053	0.050	0.033	0.031
	Poisson	0.080	0.077	0.065	0.063	0.053	0.052
$\alpha = 0.5, r = 12$ $p = 0.8$	B. Negativa	0.073	0.078	0.053	0.058	0.031	0.034
	Poisson	0.073	0.079	0.057	0.064	0.042	0.047
$\alpha = 0.8, r = 12$ $p = 0.8$	B. Negativa	0.074	0.084	0.052	0.060	0.031	0.035
	Poisson	0.070	0.081	0.051	0.059	0.034	0.039
<i>Bhattacharyya distance</i>							
Parametri	Modello	$n=50$		$n=100$		$n=300$	
		$h = 1$	$h = 2$	$h = 1$	$h = 2$	$h = 1$	$h = 2$
$\alpha = 0.2, r = 1.3$ $p = 0.3$	B. Negativa	0.0077	0.0071	0.0037	0.0032	0.0012	0.0011
	Poisson	0.0660	0.0658	0.0635	0.0630	0.0617	0.0615
$\alpha = 0.5, r = 1.3$ $p = 0.3$	B. Negativa	0.0089	0.0101	0.0043	0.0050	0.0013	0.0015
	Poisson	0.0491	0.0545	0.0471	0.0523	0.0427	0.0480
$\alpha = 0.8, r = 1.3$ $p = 0.3$	B. Negativa	0.0086	0.0111	0.0040	0.0050	0.0013	0.0017
	Poisson	0.0314	0.0377	0.0300	0.0357	0.0276	0.0322
$\alpha = 0.2, r = 12$ $p = 0.8$	B. Negativa	0.0064	0.0055	0.0037	0.0030	0.0012	0.0010
	Poisson	0.0073	0.0064	0.0050	0.0043	0.0029	0.0029
$\alpha = 0.5, r = 12$ $p = 0.8$	B. Negativa	0.0073	0.0084	0.0036	0.0042	0.0012	0.0014
	Poisson	0.0070	0.0083	0.0040	0.0047	0.0020	0.0024
$\alpha = 0.8, r = 12$ $p = 0.8$	B. Negativa	0.0064	0.0085	0.0030	0.0039	0.0011	0.0015
	Poisson	0.0060	0.0083	0.0031	0.0042	0.0015	0.0018

Capitolo 4

Applicazione su dati reali

4.1 La serie storica

In questo capitolo sono applicate le procedure di stima e previsione descritte in precedenza, per modellare una serie storica di dati reali. Lo scopo è di confrontare diversi modelli per capire quale possa avere il miglior adattamento.

La serie storica in questione riguarda l'ambito medico e, in particolare, rappresenta il numero mensile di nuove infezioni di Poliomielite negli USA. Le osservazioni disponibili vanno da Gennaio 1970 a Dicembre 1987, la numerosità campionaria è quindi pari a 216. I valori assunti dalla serie sono compresi tra un minimo di 0 e un massimo di 14.

Per capire se sia sensato applicare un modello INAR(1) su una serie di conteggio è utile verificare che le funzioni di autocorrelazione parziale e totale della serie storica siano coerenti con quelle del processo teorico. Come si nota nella figura 4.1, l'autocorrelazione totale decresce esponenzialmente all'aumentare del ritardo temporale e l'autocorrelazione parziale è significativa soltanto al primo ritardo. Ciò è perfettamente compatibile con un processo INAR(1).

La media campionaria è pari a 1.12, mentre la varianza campionaria è pari a 2.98; sembra quindi evidente una situazione di sovradisersione.

Le analisi preliminari portano a pensare che la specificazione più adatta do-

rebbe essere il modello $\text{INAR}(1)$ con distribuzione Binomiale Negativa.

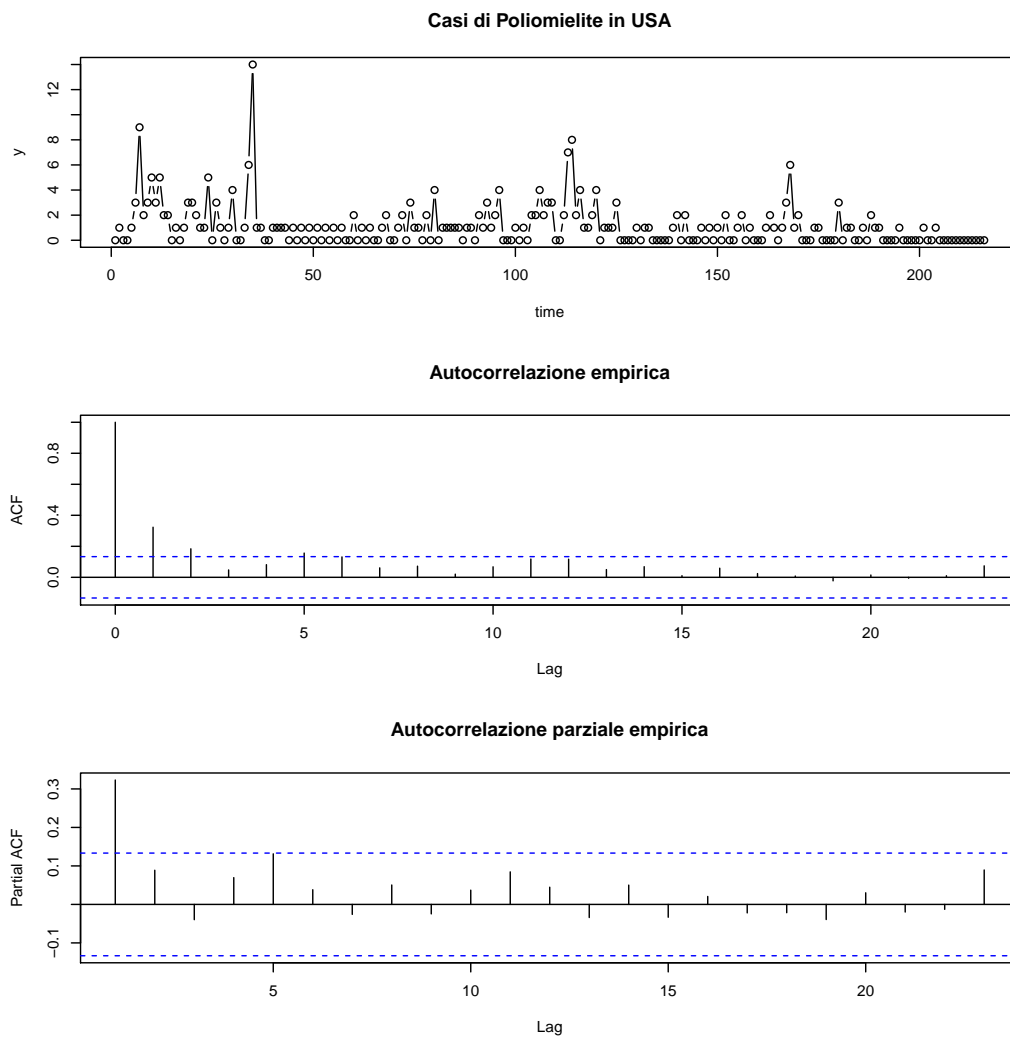


Figura 4.1: Rappresentazione grafica della serie storica e delle funzioni di autocorrelazione totale e parziale empiriche.

4.2 Confronto tra modelli

Dalle analisi preliminari l'evidente sovradisersione suggerisce l'utilizzo della Binomiale Negativa e, come si è visto nello studio di simulazione, questa dovrebbe portare a dei vantaggi in termini di previsione della distribuzione di probabilità h passi in avanti.

In questa sezione si cerca di capire quale tra il modello PoINAR(1) e il modello INAR(1) con errore Binomiale Negativa abbia un miglior adattamento alla serie.

I due modelli sono confrontati sulla base di tre diversi criteri. Il secondo e il terzo criterio sono adottati come conseguenza del fatto che non è possibile valutare direttamente la bontà delle previsioni della distribuzione di probabilità h passi in avanti, dato che sui dati reali il processo generatore è ignoto e non è possibile quindi confrontare direttamente le distribuzioni condizionali previste con quelle vere. Il primo criterio utilizza i tre indicatori, visti nel precedente capitolo, per valutare le previsioni puntuali delle ultime 100 osservazioni della serie. I risultati sono esposti in tabella 4.1.

Il secondo criterio utilizza un indicatore basato sulla probabilità che la distribuzione prevista dà al valore realizzato. Questo indicatore è calcolato con la formula: $\frac{1}{100} \sum_{t=115}^{214} \hat{p}(X_{t+h} = x_{t+h})$, per $h=1,2$ e dove x_{t+h} è l'osservazione al tempo $t+h$. I risultati sono esposti in tabella 4.2.

Il terzo criterio si basa sul confrontare la bontà della previsione della distribuzione di probabilità h passi in avanti quando h tende a infinito. La previsione è quindi data dalla distribuzione marginale del processo. La vera distribuzione marginale del processo che ha generato la serie storica non è nota, ma può essere stimata attraverso l'istogramma di frequenza delle osservazioni. I risultati di questo confronto sono presentati per via grafica nella figura 4.2.

Tabella 4.1: Criteri di valutazione delle previsioni ad h passi per le ultime 100 osservazioni della serie storica. Indicatori utilizzati (da sx a dx): errore assoluto medio, errore quadratico medio e frequenza di previsioni corrette.

Modello	Previsione	FMAE		FMSE		Frequenza	
		$h = 1$	$h = 2$	$h = 1$	$h = 2$	$h = 1$	$h = 2$
B. Negativa	Mediana	0.79	0.81	1.17	1.13	0.32	0.28
	Media	0.86	0.89	1.20	1.25	0.25	0.23
Poisson	Mediana	0.87	0.81	1.29	1.13	0.26	0.28
	Media	0.87	0.85	1.19	1.17	0.26	0.25

Tabella 4.2: Densità, della distribuzione prevista ad h passi, data al valore realizzato. Indicatore calcolato sulle ultime 100 osservazioni del campione.

Modello	$h = 1$	$h = 2$
B. Negativa	0.337	0.329
Poisson	0.284	0.301

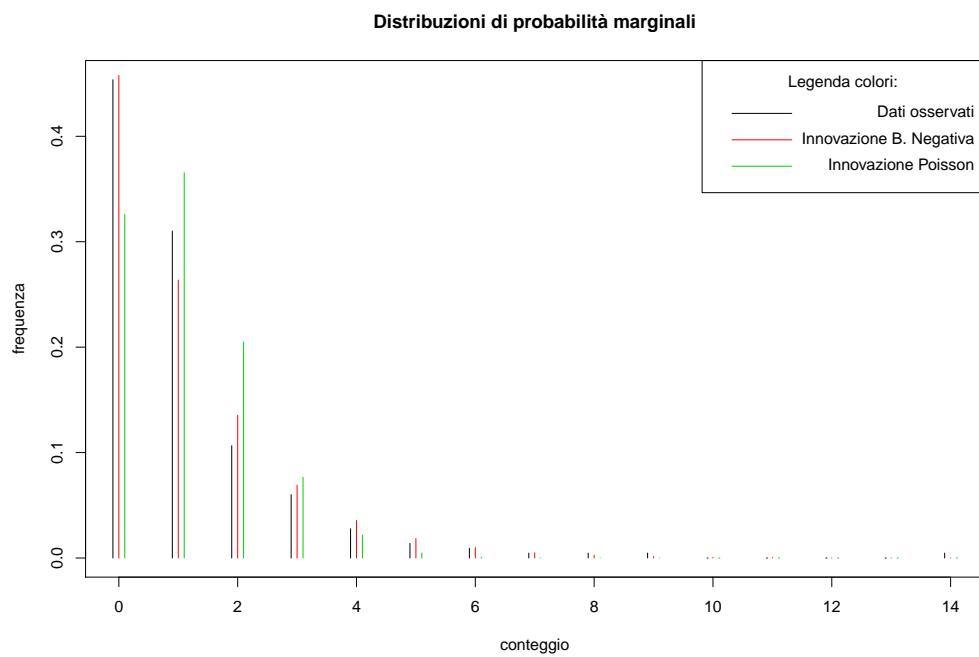


Figura 4.2: In nero: istogramma di frequenza delle osservazioni della serie. In rosso: stima della distribuzione marginale del processo quando l'innovazione è Binomiale Negativa (calcolata tramite simulazione). In verde: stima della distribuzione marginale del processo quando l'innovazione è Poisson.

I risultati portano a concludere che, il modello con la Binomiale Negativa abbia un miglior adattamento alla serie storica in questione. Questo è suggerito sia dall'indicatore della tabella 4.2 che dalla figura 4.2. Quest'ultima mostra chiaramente come la marginale stimata dal modello Binomiale Negativa sia molto più adatta rispetto a quella Poisson a rappresentare la distribuzione delle osservazioni.

I risultati di questa applicazione su dati reali rafforzano quanto evidenziato dallo studio di simulazione.

Conclusioni

Nei casi in cui l'analisi della serie di conteggio ha come unico scopo quello di ottenere delle semplici previsioni puntuali, da un punto di vista pratico, il classico AR(1) può essere considerato il modello più adatto. Questo sulla base della semplicità di applicazione e la qualità delle *performance* previsive evidenziata dallo studio di simulazione. Sebbene sia una soluzione poco elegante, le previsioni coerenti con la caratteristica di conteggio delle osservazioni possono essere ricavate semplicemente arrotondando la media condizionale all'intero più vicino. Chiaramente, non sempre le previsioni puntuali sono l'unico obiettivo e quindi, in tali situazioni, i modelli INAR(1) risultano preferibili rispetto all'AR(1). Gli INAR(1) consentono di prevedere probabilità condizionali h passi in avanti grazie ad una più adeguata rappresentazione del processo di conteggio generatore dei dati.

Gli esperimenti effettuati hanno evidenziato che, nei casi in cui sia fortemente implausibile l'ipotesi di equidispersione della serie storica, l'utilizzo del modello PoINAR(1) porta a scarsi risultati nella previsione della distribuzione di probabilità h passi in avanti. In tale senso i modelli INAR(1) con errore Binomiale e Binomiale Negativa sono da considerarsi alternative migliori rispetto al Poisson nei casi rispettivamente di sottodispersione e sovradispersione.

In questa tesi sono state trascurate alcune tematiche riguardanti le tecniche di stima e i risultati asintotici degli stimatori. Sviluppi successivi potrebbero mirare a ottenere le distribuzioni asintotiche degli stimatori di massima verosimiglianza e testare altre tecniche di stima dei parametri. Si deve inoltre

tener presente che, i due modelli proposti non sono da considerarsi come le migliori alternative al Poisson nei casi di sovra e sotto dispersione. Questi, infatti, possono presentare alcune limitazioni; ad esempio, il modello Binomiale ha il problema di produrre un supporto limitato che in molti casi può non risultare sensato. In tal proposito un ulteriore sviluppo potrebbe essere quello di considerare distribuzioni diverse da quelle trattate in questa tesi, così da verificare se ci siano soluzioni migliori in termini di flessibilità e adattabilità.

Bibliografia

- [1] M. A. Al-Osh and A. A. Alzaid (1987). *First-Order Integer-Valued Autoregressive (INAR(1)) Process*. Journal of Time Series Analysis, 8, 261-275
- [2] M. A. Al-Osh and A. A. Alzaid (1990). *An integer-valued p th-order autoregressive structure (INAR(p)) process*. J. Appl. Probab., 27, 314-324.
- [3] A. Bhattacharyya (1986). *On a measure of divergence between two statistical populations defined by their probability distributions*. Bulletin of the Calcutta Mathematical Society, 35, 99-109.
- [4] K. Fokianos (2012). *Count Time Series Models*. In Handbook of Statistics Vol 30: Time Series - Methods and Applications, eds. T. Subba Rao, S. Subba Rao and C. R. Rao, Elsevier B.V., Amsterdam, 315-347.
- [5] R.K. Freeland (1998). *Statistical analysis of discrete time series with application to the analysis of Worker's Compensation Claims Data*. Ph.D. Thesis, The University of British Columbia, Canada.
- [6] R.K. Freeland, B.P.M McCabe (2004). *Forecasting discrete valued low count time series*. International Journal of Forecasting, 20, 427-434.
- [7] G. K. Grunwald, R. J. Hyndman, L. Tedesco and R. L. Tweedie (2000). *Non-Gaussian conditional linear AR(1) models*. Australian and New Zealand Journal of Statistics, 42, 478-495.

- [8] K. Jayakumar (1995). *An integer valued autoregressive (INAR(p)) time series model*. Bulletin of the International Statistical Institute, Proceedings of the 50th Session Book, 3, 1365-1374.
- [9] B.P.M. McCabe and G.M. Martin (2005). *Bayesian of low count time series*. International Journal of Forecasting, 21, 315-330.
- [10] E. McKenzie (1985). *Some simple models for discrete variate time series*. Water Resources Bulletin, 21, 645-650.
- [11] E. McKenzie (1986). *Autoregressive moving-average processes with negative-binomial and geometric marginal distributions*. Advances in Applied Probability, 18, 679-705.
- [12] J. A. Nelder and R. Mead (1965). *A simplex method for function minimization*. Computer Journal, 7, 308-313.
- [13] R Core Team (2013). *R: A Language and Environment for Statistical Computing*. R Foundation for Statistical Computing, Vienna, Austria.
- [14] N. Silva, I. Pereira and M. E. Silva (2009). *Forecasting in INAR(1) model*. Statistical Journal, 7, 119-134.
- [15] F. W. Stuetal and K. Van Harn (1979). *Discrete analogues of self-decomposability and stability*. Ann. Prob., 7, 893-900.